# JOURNAL OF THEORETICAL BIOLOGY

ACADEMIC PRESS

VOLUME 1, NUMBER 2 APRIL 1961

London and New York

# *Journal of Theoretical Biology*







Volume 1, 1961: 12*s*. 6*d*. ($17.00)

Private subscriptions only: 86*s*. ($12.00)

Subscription orders should be sent to Academic Press Inc. (London) Ltd., 17 Old Queen Street, London, S.W.1, except those from Canada, Central and South America, and the U.S.A. which should be sent to Academic Press Inc., 111 Fifth Avenue, New York 3, N.Y.

*J. Theoret. Biol.* (1961) **2**, 107–119

# The Biological Function of Deoxyribonucleic Acid†

CARL C. LINDEGREN

*Biological Research Laboratory, Southern Illinois University, Carbondale, Illinois, U.S.A.*

(*Received* 11 *July* 1960)

## Introduction

A current theory of genetics assumes that deoxyribonucleic acid comprises the gene, and that the different genes are discriminated from each other by the seriation of the nucleotides along the long axis of the DNA molecule. The principal basis for this theory is the fact that treatment with a DNA preparation isolated from a normal culture can transform a culture carrying a defective gene into one carrying the normal gene. It has been inferred, therefore, that DNA comprises the gene and that the different genes are different from each other because of the different arrangements of nucleotides within them. This inference has been supported by the fact that destruction of DNA by a specific enzyme destroys the capacity of a DNA preparation to effect transformation of a defective into a normal gene.

Transformation is a directed mutation of a specific defective gene into the corresponding normal gene by a chemical substance. The transformation experiments with DNA preparations (which were previously assumed to have been achieved by a purified soluble capsular carbohydrate) are now assumed to be achieved by treatment with the DNA isolated from the organism carrying the dominant gene. It has been inferred as a result of these experiments that DNA, and DNA alone, is the transforming agent.

In studies of the melezitose locus of *Saccharomyces*, transformation is achieved without DNA by the use of maltose and melezitose. In previous experiments transformations of a defective galactose locus had been achieved by the use of galactose. It has been inferred as a result of the DNA transformations that DNA carries some "genetic information". If this inference is justified we may infer that the sugars, maltose, melezitose and galactose also carry "genetic information" (Lindegren & Pittman, 1959).

The hypothesis that the gene is composed exclusively of DNA does not account for the facts: (1) that transformations can be achieved by the use

† This work has been aided by a grant from the American Cancer Society.

of saccharides without DNA, (2) that most of the chromatin in the cell (possibly as much as 90 per cent) is heterochromatin which does not carry genes (the large heterochromatinic Y-chromosome in Drosophila is nearly gene-free), and (3) that euchromatin does not always carry genes. (The B-chromosome in maize contains both euchromatin and heterochromatin but does not carry genes.)

As a substitute for the theory that DNA alone comprises the gene, I propose (1) that DNA has the specific function of controlling homologous synapsis of chromosomes, (2) that the genes are located *on the surface* of the DNA, (3) that in spite of their fundamental stability genes are to some extent malleable and can be remodelled and (4) that the transformations achieved by maltose, melezitose and galactose are effected by remodelling the surface of the gene. If genes can be remodelled by substrate, it seems reasonable to suppose that genes act as receptors of substrate to initiate the synthesis of enzymes (or as receptors of gene-products to terminate enzyme synthesis).

The discrimination between euchromatin and heterochromatin has been studied by Jack Schultz (1947) and I shall summarize his views briefly: Euchromatin is the region of the chromosome which carries the Mendelian genes; heterochromatin does not carry Mendelian genes except for a few which might be supposed to have been intercalated into the heterochromatin by translocation. Euchromatin is much poorer in deoxyribonucleic acid than heterochromatin; heterochromatin tends to remain compact and heavily staining after telophase when the euchromatin has spun out into diffusely staining uncoiled threads. The heterochromatin has been called "prochromosomes" or "chromocentres" and said to be "heteropycnotic" or "pycnotic" indicating its denser character. It has been inferred that there are two kinds of heterochromatin: the $\alpha$-type which does not uncoil and the $\beta$-type which uncoils into a diffuse mesh visible in the chromocentres of the giant salivary and nerve cell chromosomes.

The euchromatinic regions of the giant salivary chromosomes of *Drosophila* are divided into bands and interband regions (Fig. 1). The heavy bands are richer in DNA than the interbands and contain both RNA and lower polymers of DNA. Although the proteins are difficult to distinguish in the bands, the presence of arginine in the bands indicates that a basic protein is present there. The interbands carry little nucleic acid but consist mainly of proteins containing aromatic amino acids. One type of protein in the interband stains with fast green and is easily soluble in slightly acid solutions while the other is (1) unstaining, (2) resistant to acid and (3) highly elastic. These two proteins may be the histone- and tryptophane-proteins of Mirsky, Pollister & Ris (1946). In most hetero-

chromatin the interbands are absent (although the Y-chromosome in *D. buscki* (Krivshenko, 1950) has faint bands and interbands) and the DNA "capsules" are connected to each other by strands of the same kind of nucleoprotein of which they are themselves composed; in euchromatin

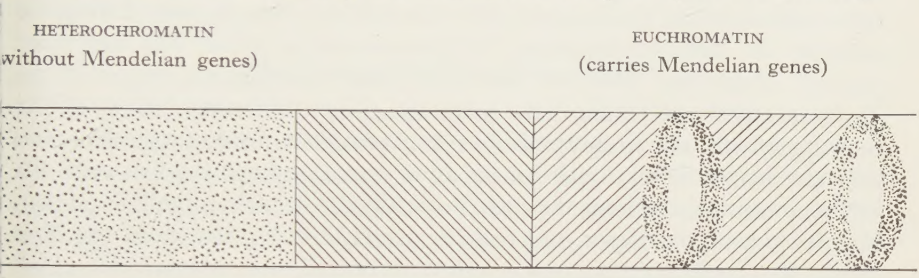



FIG. 1. Diagrammatic representation of the chromosome containing euchromatin, heterochromatin and the nucleolar connecting region.

the bands (or capsules), which are rich in DNA, are connected by interband strands of elastic protein.

The heterochromatinic regions of the different chromosomes of *D. melanogaster* all coalesce into an apparently homogeneous association at meiosis, although this does not occur in maize nor in most of the other Diptera (White, 1954). In all species the chromomeres of the euchromatin synapse in pairs, *at meiosis*, homologous chromomere with homologous chromomere, excluding the third chromonema in triploids. (Although the third chromonema is excluded *at meiosis*, all three chromosomes fuse in the salivaries of triploid *Drosophila*.) The precise pairing of euchromatinic chromomeres and the not-so-precise pairing of heterochromatin or heterochromatinic knobs indicate that each euchromatinic chromomere is unique and tends to associate in synapsis with its homologous chromomere but that the chromomeres of heterochromatin are not so distinctive, at least with regard to synapsis, for they occasionally associate in a non-homologous manner.

## The Function of Heterochromatin

Schultz has summarized the current concept of the function of heterochromatin and has made some inferences of his own: A male *Drosophila* without the heterochromatinic Y-chromosome is sterile but otherwise normal. Males that have lost the major heterochromatinic region of the X-chromosome are fertile if they contain a Y-chromosome. Many of the rearrangements on X-chromosomes in *Drosophila* produce viable males that are sterile. When both breaks in a rearrangement involve only the heterochromatinic regions the male is usually fertile; when the X is broken in a euchromatinic region the male is usually fertile, but when the X is broken in a euchromatinic region and translocated to an autosomal euchromatinic region the males are sterile. From this Schultz infers that the heterochromatins in different regions have similar functions. The heterochromatin of the Y-chromosomes is essential for normal spermatogenesis.

Schultz points out that rearrangements between euchromatin and heterochromatin lead to variegation of the characters controlled by the genes which are brought into the proximity of heterochromatin, indicating that heterochromatin changes the stability of the gene. With regard to the effect of heterochromatin on variegation, Schultz says, "The essential fact is that the regions containing the genes for which variegation is displayed show regular changes that are correlated in their frequency and extent with the frequency and extent of the variegation." This inference is supported by data from yeast which indicate that the internal chromosomal environment, especially as it concerns the availability of adenine, affects the stability of the gene (Lindegren & Lindegren, 1956). Schultz points out that in the absence of the Y-chromosome the bands in the euchromatin stain lightly and the nucleoprotein components are affected not only in amount but in arrangement as well, and he concludes that "the heterochromatic regions function in what may provisionally be termed, for lack of a more precise statement, as the generalized metabolism of the chromosome."

Schultz has inferred that heterochromatin is involved in the synthesis of nucleolar substance. In both the salivary band nuclei and the maize chromosomes the nucleolus is associated with heterochromatin. Schultz, Caspersson & Aquilonius (1940) were able to show that "the nucleoli contain ribosenucleo-proteins which were shown to vary in composition according to the genetic structure of the animal. This is the cardinal fact which lies at the base of the theory that the heterochromatic regions have to do with cytoplasmic synthesis in general, by way of precursors deriving from the nucleoli."

An adequate supply of heterochromatin (1) ensures fertility, (2) controls variegation and (3) ensures the synthesis of an adequate supply of nucleolar material. The nucleolar material is assumed to transmit the effects of gene action to the cytoplasm and there to produce the specific gene-effect, presumably in the form of an enzyme.

## The Biological Function of DNA

The theory that DNA comprises the genetic material and that crossing-over rearranges the nucleotides in the DNA implies that the synapsis of the chromonemata is effected by homologous nucleotide-to-nucleotide pairing of the strands of DNA. It is difficult to understand, however, (1) how one double helix could attract another, (2) why the attraction should be exerted only at meiosis, and (3) why only single strands should attract only other single strands, i.e. why pairing is always in two's and multiple synapses do not occur in euchromatin. I shall propose (1) that DNA does not comprise the genic material, (2) but that DNA forms a part of the chromosome on which the genes are carried, (3) that synapsis is effected by the "zippering-up" of the bases of homologously "coded" single helices in a narrow band on one side of the euchromatinic regions of the chromosome at zygotene, (4) that the alignment of nucleotides is relatively permanent and not altered by crossing-over or mutation, but (5) that the alignment serves to achieve precisely homologous synapsis of the euchromatin in which the genes are carried and thus preserves the order of the genes by assuring that crossing-over will almost always be precisely reciprocal thus avoiding the production of either deletions or duplications in euchromatin.

The 2-by-2 pairing of homologous chromonemata was demonstrated by Newton & Darlington (1929) in a study of the synapsis of triploids. They found that, when three strands synapsed at leptotene, only 2 of the 3 chromonemata were paired at any given point. The exclusion of the third chromonema in the pairing of trisomics and triploids indicates that each chromonema has only a single pairing valence or pairs only on one side. Darlington supposed that the 2-by-2 pairing was an electrical phenomenon but this fails to account for its predominantly homologous nature, Lindegren & Bridges (1938) proposed that synapsis might be explained by the possible antigenic nature of the chromomeres but this might be expected to lead to multiple associations.

Euchromatin and heterochromatin synapse quite differently. The gene-carrying euchromatin of normal chromosomes and the gene-free euchromatin of the B-chromosome both undergo precisely homologous pairing. In contrast with euchromatin (either gene-carrying or gene-free), the heterochromatin of different nonhomologous chromosomes in the salivary

chromosomes of *Drosophila* fuses in a multiply associated unaligned mass. This suggests that the DNA of heterochromatin is not so precisely arranged as that of euchromatin. The heterochromatinic regions fuse *at meiosis*, but not at other times, suggesting that the DNA of heterochromatin is present in single helices at meiosis. The capacity of DNA for effecting precise homologous pairing of the gene-bearing euchromatin ensures precisely reciprocal crossing-over, thus maintaining the genome completely intact while permitting exchanges between different sections of the gene-bearing euchromatin.

## The B-Chromosome of Maize

McClintock (1933) studied the synapsis of trisomic B-chromosomes in maize. Figure 2 is a reproduction of six of her figures describing the association at meiosis of three B-type chromosomes in a maize plant. Concerning the B-type chromosome, Rhoades and McClintock (1935) state: "There has been found, however, in certain strains of maize, especially Black Mexican sweet corn, a type of supernumerary chromosome which is totally unlike any members of the regular complement. It cannot be said, on the basis of its peculiar morphology, to have been derived from any one member of the monoploid set and its origin is unknown. This type of supernumerary has been called the B-type chromosome in contradistinction to members of the normal complement which Randolph (1928) has designated as the A-type chromosome. The B-type has a distinctive morphology quite unlike any of the A-type chromosomes. It appears to be composed of genetically inert material and carries no known genes. The presence of one or many B-type chromosomes has no visible effect on the morphological character of the plant." (This proves that the B-chromosome does not carry genes.) "Randolph has succeeded through successive crosses in accumulating more than twenty-five B-type chromosomes in a single plant in addition to the regular complement of twenty. In contrast with supernumeraries composed of A-type chromosomes the B-type is readily transmitted through both pollen and eggs.

"The morphology of the B-type chromosome at pachytene and its synaptic behavior have been investigated by McClintock (1933). In the meiotic prophase the B-type is slightly more than one-half the length of the shortest member of the normal complement. Its morphology at midprophase, beginning with the terminal insertion region, is as follows: (1) terminal spindle fiber attachment region, (2) pycnotic (heterochromatic) region, (3) chromatic (euchromatic) region composed of small but distinct chromomeres, (4) elongate pycnotic region with several definite constrictions, (5) bulging pycnotic region, (6) broken pycnotic region composed of four distinct parts."

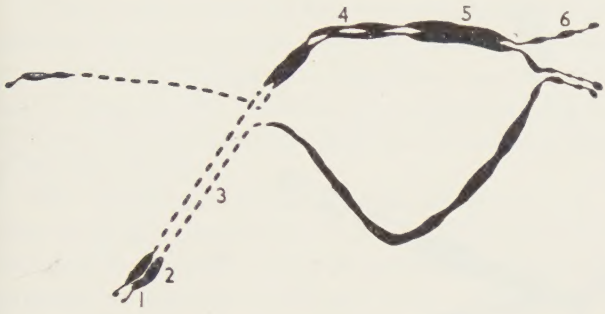



FIG. 4

FIG. 5

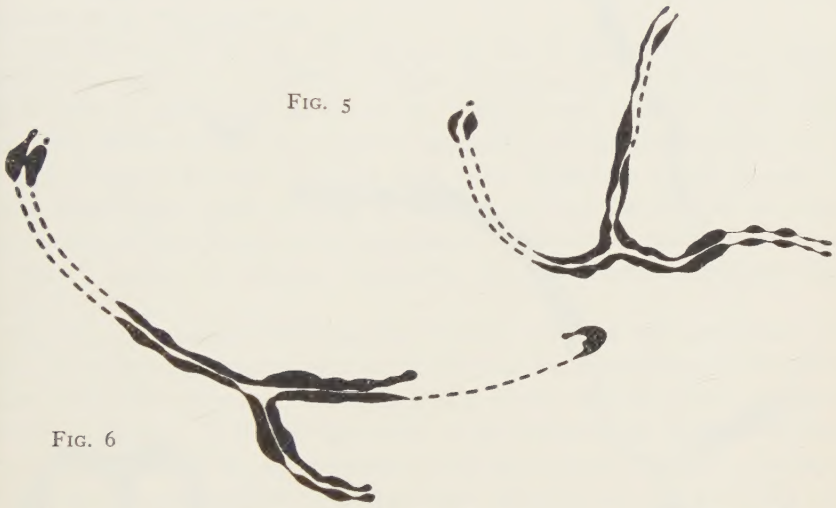


FIG. 6

FIG. 4. The association of three B-type chromosomes. In this figure only homologous association has occurred. Plant 536–2.

FIG. 5. The association of three B-type chromosomes to produce a T configuration. Homologous association produced the two side arms. The upper arm resulted from nonhomologous association. Plant S 31–314–2b.

Fig. 6. The association of three B-type chromosomes to produce an asymmetrical T configuration. The 2-by-2 association in the arm to the right is nonhomologous. Plant S 31–314–2b.

FIG. 2a. Figures 4 to 6 from McClintock's 1933 paper. Mag. × 1650. (By courtesy of Springer-Verlag, Heidelberg).

(I have numbered these six regions in McClintock's Fig. 4. The only euchromatinic region in this chromosome is region 3, which Rhoades and McClintock called a "chromatic" region. Regions 2, 4, 5 and 6 are heterochromatinic regions for which they use the older term "pycnotic".)

FIG. 7


FIG. 8

FIG. 9

FIG. 7. The association of three B-type chromosomes. Only homologous parts are associated 2-by-2. The univalent threads to the lower right and left have remained unassociated. Plant 536–2.

FIG. 8. The association of three B-type chromosomes. The chromosomes are numbered 1, 2 and 3. Chromosome 1 associated at the spindle fiber attachment end with chromosome 2. Chromosome 1 associated at the opposite end with chromosome 3. The rest of chromosome 2 is in the form of a foldback. Such figures are comparatively frequent. Plant 536–2.

FIG. 9. The association of three B-type chromosomes to form a complex configuration. Such complex trivalent configurations have been observed occasionally in plants trisomic for a member of the monoploid complement. Plant 536–2.

FIG. 2b. Figures 7 to 9 from McClintock's 1933 paper. Mag. × 1650. (By courtesy of Springer-Verlag, Heidelberg).

"The greater part of the B-type at mid-prophase is composed of pycnotic material. As stated before, there is reason to believe that the B-type chromosome is genetically inert. That these two facts have some close relationship is suggested by Heitz's studies (1933, 1934) with *Drosophila* in which he shows that the pycnotic portions found in prophase chromosomes are genetically inert.

"The B-type shows no synaptic affinity for any of the chromosomes composing the normal set. If a single B-type is present it behaves as a univalent but regions within it are often non-homologously paired at mid-prophase in meiosis. Synapsis occurs between B-type chromosomes if two or more are present in the same nucleus although non-homologous association is very common.

"McClintock found that in plants with two B-types there were more sporocytes with two univalent B-type chromosomes at diakinesis than there were in mid-prophase where the B-types were usually paired. This has been attributed to the complete separation of the two members of a B-type bivalent during diplotene and early diakinesis which occurs in some sporocytes. This precocious separation may be due to frequent occurrence of non-homologous association observed at pachytene and/or to partial failure of chiasma formation when homologous pairing does occur."

The legends under McClintock's figures describe the association of three B-type chromosomes and point out that the association is always in pairs, the third chromatid being excluded in any region where two have already synapsed. It was in the study of the synapsis of B-chromosomes that she discovered the phenomenon of nonhomologous synapsis; she says, in private communication, that it is more conspicuous but probably not more common between paired B-type than between paired normal chromosomes. If one assumes that these six figures give a good sample of the different kinds of nonhomologous synapsis, it is clear that nonhomologous synapsis of euchromatin to euchromatin is rare. For example, the nonhomologous synapsis in her Fig. 5 occurs in the upward region of the inverted T and is a nonhomologous synapsis of heterochromatin to heterochromatin, and euchromatin to heterochromatin; nonhomologous synapsis of euchromatin to euchromatin does not occur. Similarly, in her Fig. 6 the euchromatin synapses homologously but in one part of the T there is a synapsis of two nonhomologous heterochromatinic regions, together with a fold back in the small terminal heterochromatinic region designated as 2. But the fold back does not lead to homologous synapsis in the dotted euchromatinic region which remains single. In all the other figures the same pattern holds: if nonhomologous synapsis occurs it is either between two nonhomologous heterochromatinic regions or between a heterochro-

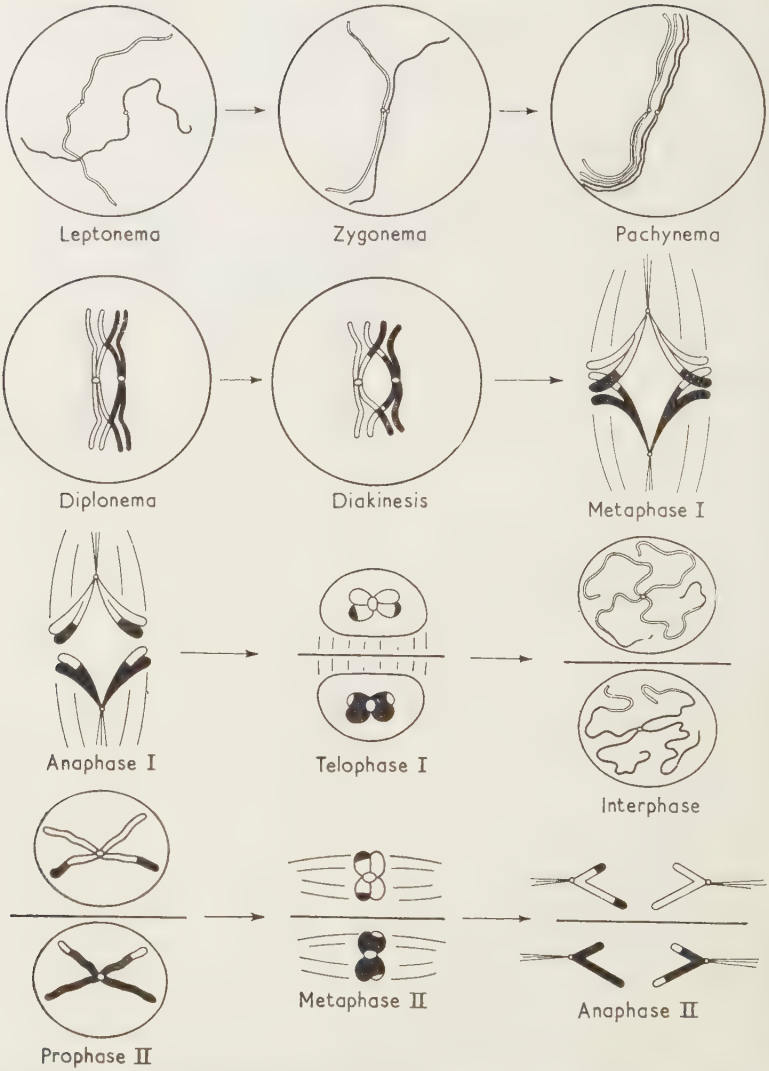



FIG. 3. The stages of meiosis. Highly schematized diagram of meiosis in which a single pair of chromosomes is followed. (Figure 1 from Rhoades's 1950 paper, by courtesy of the American Genetic Association.)

matinic and a euchromatinic region but not between two nonhomologous euchromatinic regions. On the basis of these figures, one might infer that the synapsis in a euchromatinic region is generally homologous, whether or not that region contains genes, and that synapsis between heterochromatinic regions may be either homologous or nonhomologous with a general preference for homologous synapsis. It may be inferred that the

coding of DNA in euchromatin ensures precise pairing but that the coding in heterochromatin does not invariably ensure precise pairing. I propose that the synapsis in euchromatin is due to the exposure of a narrow band of single helices of DNA and that it is this narrow band which ensures that only two strands are associated in any given region, while the third strand is excluded. On this proposal, one can account for both the precision of the pairing and the fact that only two strands are involved at any

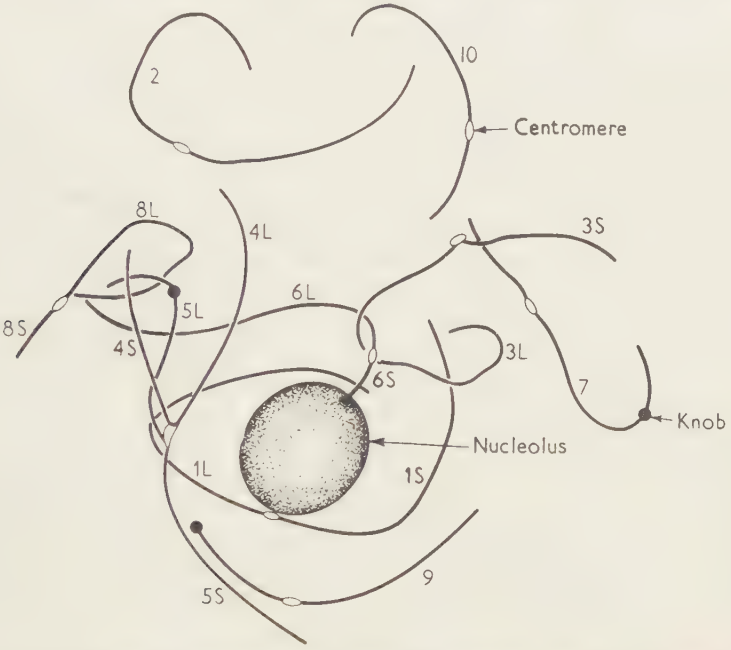



FIG. 4. The ten maize chromosomes. Schematic diagram of pachytene chromosomes: each bivalent chromosome is represented by a single line in order to simplify the diagram. (Figure 2 from Rhoades's 1950 paper, by courtesy of the American Genetic Association.)

point. One function of heterochromatin, in addition to those proposed by Schultz, may be to supply an adequate reserve of nucleotides to prevent the occurrence of deficiencies in euchromatin and thus to preserve the continuity of the genome. If the reserve nucleotides were present in short chains of open helices, this would explain the tendency of heterochromatin for nonhomologous synapsis. The nonhomologous synapsis of heterochromatin in the salivary chromosomes of Drosophila is much more pronounced than in maize; since in Drosophila all the heterochromatin fuses into a single chromocenter suggesting that most of the heterochromatin is in the form of short single helices.

The asynaptic and the sticky gene in maize studied by Beadle (1932)

disturb normal pairing at meiosis. In a general summary of maize genetics Rhoades & McClintock (1935) state that the asynaptic gene "disturbs in some way the normal pairing of chromosomes in early prophase so that most of the chromosomes are present as univalents at diakinesis and metaphase. The sticky gene (*st*), as its name implies, causes the chromosomes to stick together so that at metaphase I there is a clumped mass of chromatin rather than ten independent bivalent pairs. The sticky gene apparently increases frequency of non-disjunction, gene mutation and produces translocations, deficiencies and chromosome fragments."

It may be assumed that the asynaptic gene reduces the number of single helices which appear at meiosis. The sticky gene may cause the entire chromosome to be covered with exposed single helices at meiosis. One may infer that the conditions that lead to the exposure of coded DNA along the side of the chromosomes in a single narrow band at meiosis and not in mitosis are the result of some balanced physiological mechanism. Rhoades (1950) has described meiosis in maize (Fig. 3) and pointed out that the centromeres at meiosis often synapse nonhomologously.

"The centromeres of non-homologous chromosomes are often observed paired or stuck together at pachynema (Fig. 4) although they become free at late diplonema. Likewise the heterochromatic knobs on different chromosomes have a striking tendency to unite at pachynema. Presumably both the pairing of centromeres and knobs is due to the non-specificity of these regions, i.e. the centromere regions of the various chromosomes are of similar composition, as are the knobs (heterochromatin)." This fact may be the basis for the phenomenon of reverse linkage observed in Saccharomyces (Shult & Lindegren, 1956) or of the phenomenon of affinity observed in mice (Wallace, 1958).

## Conclusions

It is inferred that the coding of DNA ensures precise pairing of homologous regions at meiosis and that the double helix ensures precise duplication of the coded structure and that this is the biological function of DNA.

If coding of nucleotides were involved in gene action, one would expect the chromatin of the B-chromosome (which is presumed to be coded because it undergoes homologous synapsis) to produce enzyme, or to stimulate gene actions which would interfere with the gene action of the standard genome. But neither B-chromosomes nor heterochromatin "interferes" with the action of the standard genome suggesting that neither the disordered (short chain?) chromatin of the heterochromatin nor the regularly arranged gene-free nucleotides of the B-chromosome has any effect on the genes carried in the euchromatin of the standard chromosomes and that these materials have survival value in producing an

abundant reserve of short chain nucleotides which ensures that the euchromatin can always be maintained intact.

The phenomenon of position effect suggests that the genes (in those organisms that show this phenomenon) function most effectively when they are located in specific places on the chromosome. The effectiveness of DNA in transformation, transduction and the peculiar type of partial hybridization, described by Wollman, Jacob & Hayes (1956), and Jacob & Wollman (1958), are explicable if one assumes that the coded DNA of the donor's transforming principle synapses with similarly coded DNA of the recipient chromosome, inserting the chromosomal segment into a functional position in the recipient chromosome. Treatment of transforming principle with DNase would destroy the coded carrier and make the transforming principle ineffective.

On this theory, reciprocal crossing-over would not rearrange the nucleotides. (Unequal crossing-over would make the chromosome longer.) If crossing-over occurred at Ca bands between the gene-carrying segments of DNA, as proposed by Steffensen (1953; 1955), the genes would be reassorted by crossing-over, but the aperiodic arrangement of the nucleotides in a species would be permanent and species-specific.

## REFERENCES


Beadle, G. W. (1932). *Z. indukt. Abstamm.-u VererbLehre*, **63**, 195.
Heitz, E. (1933). *Z. Zellforsch.* **19**, 720.
Heitz, E. (1934). *Z. Zellforsch.* **20**, 237.
Hotchkiss, R. D. (1957). In "The Chemical Basis of Heredity", 321–335. (W. D. McElroy and B. Glass, Eds.) Johns Hopkins Press, Baltimore.
Jacob, F. & Wollman, E. L. (1958). In "The Biological Replication of Macromolecules" 75–92. *Symp. Soc. exp. Biol.* Academic Press, New York.
Krivshenko, J. D. (1950). *Proc. nat. Acad. Sci. Wash.* **36**, 703.
Lindegren, C. C. & Bridges, C. B. (1938). *Science*, **87**, 2266.
Lindegren, C. C. & Lindegren, G. (1956). *Nature, Lond.* **178**, 796.
Lindegren, C. C. & Pittman, D. (1959). *Genetica*, **30**, 169.
McClintock, B. (1933). *Z. Zellforsch.* **19**, 191.
Mirsky, A. E., Pollister, A. W. & Ris, Hans (1946). *Rec. Genet. Soc. Amer.* **14**, 54.
Newton, W. C. F. & Darlington, C. D. (1929). *J. Genet.* **21**, 1.
Randolph, L. F. (1928). *Anat. Rec.* **41**, 102.
Rhoades, M. M. (1950). *J. Hered.* **41**, 58.
Rhoades, M. M. & McClintock, B. (1935). *Bot. Rev.* **1**, 292.
Schultz, Jack (1947). *Cold. Spr. Harb. Symp. quant. Biol.* **12**, 179.
Schultz, Jack, Caspersson, J. & Aquilonius, L. (1940). *Proc. nat. Acad. Sci. Wash.* **26**, 515.
Shult, E. E. & Lindegren, C. C. (1956). *Genetics*, **28**, 165.
Steffensen, D. (1953). *Proc. nat. Acad. Sci. Wash.* **39**, 613.
Steffensen, D. (1955). *Proc. nat. Acad. Sci. Wash.* **41**, 155.
Wallace, M. E. (1958). *Phil. Trans.* **B241**, 211.
White, M. J. D. (1954). "Animal Cytology and Evolution". Cambridge University Press, London.
Wollman, E. L., Jacob, F. & Hayes, W. (1956). *Cold Spr. Harb. Symp. quant. Biol.* **21**, 141.

*J. Theoret. Biol.* (1961) **2**, 120–140

# Tracer Dynamics: I. A Tentative Approach and Definition of Fundamental Concepts

PER-ERIK E. BERGNER

*Department of Medical Physics, Karolinska Institutet, Stockholm, Sweden*

(*Received* 20 *September* 1960)

The tracer system is discussed in terms of two disjoint sets of atoms: the set of tracer atoms and the set of naturally occurring atoms—the mother substance. The two kinds of atoms have the same atomic number and differ only in respect of the atomic weight, i.e. different isotopes of one and the same element. In terms of elementary set-theoretical concepts two postulates are formulated which together define a stochastic system. As a special case the "finite" tracer system is defined as a finite Markov chain, and it is shown that this approach gives results equivalent to those usually derived from deterministic models.

The methods applied make possible a rigorous definition of concepts such as a general compartment and a compartment with "complete mixing" (here called a "subsystem"). Special attention is devoted to the concept of turnover, which is given a rigorous definition applicable to a large class of compartments.

## 1. Introduction

If we have a system in which a certain substance, in the following referred to as *the mother substance*, takes part in different processes such as diffusions and chemical reactions, we often want information about these processes. It seems then natural to assume that one could obtain such information by observing the behavior of another, easily observable substance, which is introduced in small amounts into the system, and the kinetic behavior of which can be regarded as equivalent to that of the mother substance. Such a substance is called a *tracer*. The analysis of such observational data is performed under the assumption that the amount of tracer introduced is so small that the "natural state" of the system is not disturbed.

Several papers discuss the theory of such methods (for instance, Solomon, 1949; Sheppard & Householder, 1951; Rescigno, 1954, 1956 and 1960; Berman & Schoenfeld, 1956; Hart, 1955, 1957, 1958 and 1960), but hitherto the main interest seems to have been focused on the purely

mathematical problems connected with the analysis of the tracer data.† However, such treatments must always be based on certain assumptions or postulates in order to render a formal analysis possible. One of the purposes of the present investigation, of which this paper is the first of a series, is to give a rigorous formulation of those postulates which seem necessary for the theoretical approach. Now, we may use postulates mainly in two ways:

1. We have a mathematical model which describes the observed behavior of a physical system and we postulate certain physical significances of the parameters that appear in that model, an approach which is used, for instance, in wave mechanics.

2. We postulate certain properties of the physical system and deduce from them the "expected" behavior of the system, e.g. as in statistical mechanics.

It is this last type of approach which will be used here and we may regard the postulates as a definition of a special system constituting the physical background of our formal treatment.

Postulates, when used in this way, must be reasonable from a physical point of view and must permit a formal analysis. Evidently the choice of such postulates is, within certain limits, a matter of taste. Thus, the stochastic approach formulated in the next section will later be used in a deterministic form, and it may therefore be asked if it would not have been more justified to work throughout from a deterministic point of view. However, the stochastic model seems more satisfactory, especially since we shall try to treat a somewhat more general system than the one usually considered. Further, the stochastic model has another advantage, of great importance at least from a theoretical point of view, which is that it takes into account the possibilities of random fluctuations in the behavior of the system. Even if at present such fluctuations do not seem to be of much practical interest, with experimental improvements they may one day appear important also to the experimentalist.

In this investigation we shall try to look at the system as a physical one in its own right and not so much as a purely mathematical problem. An attempt is made to put together some new results with those hitherto scattered in literature. On the other hand, as the treatments are formal and somewhat abstract, the results may sometimes look quite trivial from a purely experimental-physical point of view. However, there are hardly any definitions, assumptions or statements which are so trivial that they are not worth a rigorous formulation and, if possible, also a formal proof.

† A critical paper on a common method in this field has been published by Bergner (1959), and an error in that paper has given rise to certain discussions (Berman & Schoenfeld, 1960, and Bergner, 1960).

Often this will simplify the view and the understanding of the problems, experimental or theoretical, and it is also from this point of view that the present work has been planned.

## 2. List of Symbols

(The numbers within brackets indicate the page where the symbol is defined.)

$u$ = an individual atom with given fixed atomic number [123]
$S$ = the general tracer system [123]
(s) = the finite tracer system [128]
$\mathbf{B}_S$ = the set of all $u$ in $S$ [123]
$B_S^0$ = the set of all naturally occurring isotopes of $u$ in $S$ = the mother substance [123]
$B_S$ = the set of tracer isotopes of $u$ in $S$, not contained in $B_S^0$ [124]
$\omega$ = the non-geometrical state variable [125]
$\bar{r}$ = the geometrical state, i.e. the position vector in the three dimensional Euclidian space [124]
$\bar{q} \equiv (\omega, \bar{r})$ = the position vector in $Q$ [125]
$Q$ = the four-dimensional state space of $u \epsilon \mathbf{B}_S$ [125]
$e$ = the environments of $S$, i.e. $e = Q - S$ [125]
$S$ and $e$ = the corresponding sets of points in $Q$ [125]
$\sigma(\bar{q}, t)$ = the content density of tracer in the point $\bar{q}$ at time $t$ [126]
$\sigma^0(\bar{q}, t)$ = the content density of mother substance in the point $\bar{q}$ at time $t$ [126]
$\mathrm{P}_i(t)$ = the probability of $u \epsilon \mathbf{B}_i$ at time $t$ [128], or of any $u \epsilon B_{(s)}$ at $t = 0$ being in $(ss)_i$ at time $t$ [131]
$C_\nu$ = the compartment $\nu$ of $S$ [125]
$\mathbf{B}_\nu = \mathbf{B}_S \,.\, C_\nu = B_\nu \cup B_\nu^0$ [126]
$b_\nu(t)$ = the content of $B_\nu$ at time $t$, in units of gram atoms [126]
$b_\nu^0(t)$ = the content of $B_\nu^0$ at time $t$, in units of gram atoms [126]
$f_\nu(\bar{q}, t)$ and $f_\nu^0(\bar{q}, t)$ = the relative content densities of $B_\nu$ and $B_\nu^0$ [126]
$(ss)_i$ = the subsystem $i$ of (s) [128]
Greek letters $\nu$ and $\mu$ are used to indicate compartments while latin letters $i$ and $j$ indicate subsystems
$P_{ij}(\Delta t)$ = the transition probability for the transition $(ss)_j \to (ss)_i$ during time $\Delta t$: (s) is assumed to be stationary [128]
$r$ = the number of (ss) of which (s) consists [128]
$\lambda_{ij}$ = the probability factor for the flux $(ss)_j \to (ss)_i$, and equal to $\dfrac{d}{d\Delta t} P_{ij}(\Delta t)|_{\Delta t = 0}$ [129]

$\zeta_{ij}(t)$ = the macroscopic flux $(ss)_j \to (ss)_i$ of tracer at time $t$ [130]
$\zeta^0_{ij}(t)$ = the macroscopic flux $(ss)_j \to (ss)_i$ of mother substance at time $t$ [130]
$Q'_\nu = Q - C_\nu$, i.e. the complement of $C_\nu$ relative to $Q$ [136]
$C'_\nu = S - C_\nu$, i.e. the complement of $C_\nu$ relative to $S$ [136]
$\mathbf{B}'_\nu = \mathbf{B}_S \cdot C'_\nu$ [136]
$\lambda(\bar{q}, \nu)$ = the turnover coefficient of $\mathbf{B}_\nu$ [136]
$\varphi_\nu(t)$ = the tracer turnover function of $\mathbf{B}_\nu$ [136]
$\varphi'_\nu(t)$ = the tracer turnover function of $\mathbf{B}'_\nu$ [136]
$A_j = b_j/b^0_j$, the specific activity in $(ss)_j$ [130]
$\varphi^0_\nu(t)$ = the turnover function of $\mathbf{B}_\nu$ [137]
$\lambda^0_\nu$ = the turnover factor of $\mathbf{B}_\nu$ [137]
$\lambda_\nu$ = the tracer turnover factor of $\mathbf{B}_\nu$ [139]
$h^0_\nu$ = the turnover of $\mathbf{B}_\nu$ [137]
A dot over a symbol indicates differentiation with respect to $t$ [129]
$\lambda$ = the square matrix of order $r \times r$ of the elements $\lambda_{ij}$ [131]
$\bar{b}(t)$ = the column vector $\begin{pmatrix} b_1(t) \\ \vdots \\ b_r(t) \end{pmatrix}$ [132]
$N$ = Avogadro's number [130]
$\zeta_{\nu e}(t)$ = the flux of tracer from $e$ into $C_\nu$ [137]

## 3. The Postulates

In the following the term *tracer system* $S$ will mean any kind of physical system on which tracer kinetic experiments can be performed; that is, the physical object of our research, such as a rabbit, or more exactly, those parts of the rabbit where the tracer will be found after being given to the animal. Therefore $S$ does not necessarily contain any tracer.

Let $u$ stand for an individual atom with *fixed atomic number* and let $\mathbf{B}_S$ stand for the set of all such $u$ in $S$. Thus $u \epsilon \mathbf{B}_S$ ($u$ belonging to $\mathbf{B}_S$) means that $u$ is somewhere in $S$ and has a given atomic number. We may now consider different isotopes $I$ of $u$, so that $\mathbf{B}_S$ can be regarded as a union of subsets

$$\mathbf{B}_S = \mathsf{U}_{I = 0, 1, 2, \ldots} B^I_S \tag{1}$$

where $B^0_S$ = the set of naturally occurring isotopes of $u \epsilon \mathbf{B}_S$, the *mother substance*, while $B^I_S \mid_{I = 1, 2, \ldots}$ = the sets of the different isotopes of $u \epsilon \mathbf{B}_S$ which do not belong to $B^0_S$ and are given as *tracer* to $S$. In the following we shall limit ourselves to the case where only one kind of isotope is used as tracer and we can therefore drop the index $I$ and write

$$\mathbf{B}_S = B_S \cup B_S^0 \tag{2}$$

where $B_S$ = the set of that $u$-isotope $\epsilon\mathbf{B}_S$ which is used as tracer for the mother substance $B_S^0$. Evidently $B_S$ and $B_S^0$ are disjoint, or

$$B_S \cap B_S^0 = \phi \tag{3}$$

where $\phi$ stands for the *empty set*, i.e. $B_S$ and $B_S^0$ have no elements in common. In other words, $B_S \cup B_S^0$ is the set of all $u$ belonging to $B_S$ and/or $B_S^0$, i.e. the set of $u\epsilon(B_S$ and/or $B_S^0)$, while $B_S \cap B_S^0$ is the intersection of $B_S$ and $B_S^0$ and equal to the set of $u\epsilon(B_S$ and $B_S^0)$, but as there are no such $u$ this set is empty.

The concept of set is here used in a purely physical sense. Therefore, if a set is empty (i.e. $= \phi$) then it means that an element belonging to that set is a physical impossibility. Now, if the atomic number of $u$ is 53 then $B_S^0$ is the set of all $^{127}$J in $S$, and if the tracer is chosen to be $^{130}$J then $B_S$ is the set of all $^{130}$J in $S$, even before a tracer is given to $S$. Thus, a set may be a purely conceptual thing and it is non-empty when it has a physical sense; we may speak of the set of $^{130}$J in $S$ even if there are no such isotopes in $S$ at present, in which case we say that the *content* of $B_S$ is zero. The content of a set is here defined as the number of elements belonging to it. It is also true that the content of $\phi$ is always zero, a zero content of a particular set not necessarily implying that this set is empty; it makes sense to think of the elements also when they are not present.

Now, $u\epsilon\mathbf{B}_S$ can be found in different geometrical positions $\bar{r}$, where $\bar{r}$ is the geometrical position vector in the three-dimensional Euclidian space, and we shall refer to $\bar{r}$ as the *geometrical state* of $u$. However, the state of $u$ is of course not completely described by $\bar{r}$, since at a certain point $\bar{r}$ there will often, during a finite time interval, appear several $u\epsilon\mathbf{B}_S$ in different physical states, e.g. $u$ may be a constituent of different chemical components the molecules of which are moving around in $S$. As a somewhat more specific example we may take the potassium ions in blood, which can be found in the plasma or in the red blood cells, and within the red cells there are probably different macro-molecules to which the potassium ions can be bound. All such states in which $u$ can be found, and which do not refer to any special geometrical position of $u$, will here be called *non-geometrical states*. As an additional example of this kind of state we have the energy content of $u$.

When we speak of a complete description of $u$ we shall here mean all the geometrical and non-geometrical coordinates which are necessary for "explaining" the macroscopic behavior of $u$. This is indeed a somewhat vague concept, but it is not catastrophic since we usually have some intuitive understanding of what we mean by the "state" of an atom—

see the examples given above. So, instead of analysing this in itself very interesting and fundamental concept, we shall, with $e$ standing for the *environments* or surroundings of $S$, simply accept the following:

*Postulate* 1. The set of all non-geometrical states, which together with the geometrical states are sufficient for a complete description of the states of $u \epsilon \mathbf{B}_S$, is finite and uniquely defined by $e$, $S$ and the atomic number of $u$.

In view of the fact that we always observe $S$ with a finite accuracy, i.e. since there are always quantities which are too small to be observable, it seems natural to assume that a finite number of non-geometrical states is sufficient for a complete description of $u$ when the geometrical states also are given. In some respects this set of sufficient non-geometrical states evidently depends on the experimental procedure, but as it is desirable to have the complete system defined independently of how the observations are performed, we have here postulated that the set is uniquely given by $e$, $S$ and the atomic number of $u$. This now means that we can define a variable $\omega$ which only takes positive integer values 1, 2, 3, . . . , $W < \infty$ corresponding to the different non-geometrical states.

Thus, according to postulate 1, the state of a $u \epsilon \mathbf{B}_S$ is completely given by the point $(\bar{r}, \omega) \equiv \bar{q}$ in the four-dimensional space $Q$. We may regard $Q$ as a *general state space* which contains all the sufficient states of $u \epsilon \mathbf{B}_S$, and it is given when $e$, $S$ and the atomic number of $u$ are given. We can then look upon $S$ as a set of points in $Q$, the complement of which relative to $Q$ is

$$Q - S = e$$

and we shall use the symbols $S$ and $e$ also for the corresponding sets. When necessary we shall explain whether $S$ and $e$ stand for the respective physical concepts or just for the sets of points in $Q$. In order to avoid any misunderstandings it should perhaps be pointed out that when we speak of all the sufficient states of $u$ we understand all the states which are of significance with respect to the behavior of $u \epsilon \mathbf{B}_S$. For instance, if $u \epsilon \mathbf{B}_S$ can leave $S$ for $e$ or if during the experiment $u$-atoms are moving from $e$ into $S$ (e.g. by injections), then certainly the states $\bar{q} \epsilon e$ are of importance and must be included in the set of sufficient states. It should be observed that while $S$ constitutes a "region" in the Euclidian space this is not always so in the $Q$-space, i.e. it is not always the case that two points $\bar{q}_1$ and $\bar{q}_2$ belonging to $S$ can be joined by a continuous curve in $S$. So, if we speak of $u \epsilon \mathbf{B}_S$ moving along certain lines in $Q$, these lines of transitions may indeed be discontinuous.

By a *compartment* $C_\nu$ of $S$ we understand a set of points in $Q$ such that $C_\nu \subset S$. To each $C_\nu$ there corresponds a subset of $\mathbf{B}_S$ containing those $u \epsilon \mathbf{B}_S$ which are in the states $\bar{q} \epsilon C_\nu$, and we shall for such a set use the notation

or
$$\left.\begin{aligned} \mathbf{B}_\nu &= \mathbf{B}_S \cdot C_\nu \\ B_\nu &= B_S \cdot C_\nu,\ B_\nu^0 = B_S^0 \cdot C_\nu \end{aligned}\right\} \tag{4}$$

In a preceding paragraph we touched upon the concept of content. We shall now define it explicitly in terms of the *content densities* $\sigma(\bar{q}, t)$ and $\sigma^0(\bar{q}, t)$ such that

$$b_\nu(t) \equiv \int_{C_\nu} \sigma(\bar{q}, t)\, d\bar{q},\ b_\nu^0(t) \equiv \int_{C_\nu} \sigma^0(\bar{q}, t)\, d\bar{q} \tag{5}$$

where $b_\nu(t)$ and $b_\nu^0(t)$ are the number at time $t$ of $u \epsilon B_\nu$ and $u \epsilon B_\nu^0$ respectively, expressed in units of gram-atoms. In the integrals $d\bar{q}$ refers to the $Q$-space, and if $dv$ refers to the ordinary volume of the three-dimensional Euclidian space we have

$$b_\nu(t) = \sum_{\omega_i \epsilon C_\nu} \int_{V_\nu} \sigma(\omega_i, \bar{r}, t)\, dv$$

where $V_\nu$ = the Euclidian region of $C_\nu$ and where the summation is taken over all $\omega_i$ belonging to $C_\nu$. It is in this way the integrals (5) should be read. From (2) and (3) it then follows that the total content of $\mathbf{B}_S$ is

$$b_S(t) + b_S^0(t) = \int_S \{\sigma(\bar{q}, t) + \sigma^0(\bar{q}, t)\}\, d\bar{q} \tag{6}$$

We shall also introduce the *relative content densities* of $B_\nu$ and $B_\nu^0$ defined by

$$f_\nu(\bar{q}, t) \equiv \frac{\sigma(\bar{q}, t)}{b_\nu(t)},\ f_\nu^0(\bar{q}, t) \equiv \frac{\sigma^0(\bar{q}, t)}{b_\nu^0(t)} \tag{7}$$

Since from (5) we have

$$\int_{C_\nu} f_\nu(\bar{q}, t)\, d\bar{q} \equiv \int_{C_\nu} f_\nu^0(\bar{q}, t)\, d\bar{q} \equiv 1 \tag{8}$$

these two functions define the distribution of the tracer and its mother substance in $C_\nu$.

In the following we shall always assume that

$$\sigma(\bar{q}, t)\Big|_{\substack{\bar{q} \epsilon S \\ t \geqslant 0}} \ll \sigma^0(\bar{q}, t)\Big|_{\substack{\bar{q} \epsilon S \\ t \geqslant 0}} \tag{9}$$

and, throughout, all discussions of $S$ will be performed under the assumption of (9) and also that $t \geqslant 0$. Now, if we regard a $u \epsilon \mathbf{B}_S$ in a certain state $\bar{q}_j$ at time $t$ (which is thus greater than or equal to zero), we may speak of the probability that this particular $u$ during time $\Delta t$ shall move to the state $\bar{q}_i$, i.e. be in $\bar{q}_i$ at $t + \Delta t$.

*Postulate* 2. The probability of an arbitrary transition $\bar{q}_j \to \bar{q}_i$ in $Q$ during $\Delta t$ at time $t$ is the same for all $u \epsilon \mathbf{B}_S$ independent of their prior histories and also independent of $b_S(t)$ and the distribution of tracer in $Q$.

These transition probabilities will be referred to as the *dynamic properties of u*. By the *dynamic state of S* with respect to $u$, i.e. to a certain atomic number, we understand the set of all such probabilities and, when they are all independent of $t$, $S$ is said to be *stationary*. Thus, what we have postulated is that the dynamic properties of $u$ when in a state $\bar{q}$ are independent of how it appeared in that state. This may be expressed as: on a macroscopic level the "memory" of $u$ can be neglected. We consider only the macroscopic behavior of $S$, and, in the same way as postulate 1, postulate 2 must in the first instance be looked upon as an approximation. This is also true for that part of postulate 2 which states that the different $u$-isotopes have identical dynamic properties, a problem which has been treated, for instance, by Bigeleisen (1949). As a general rule one could probably say that, except for the hydrogen isotopes, the isotopic effect could be neglected, especially when the system is close to equilibrium.

Another feature of postulate 2, which shows the approximate character of that postulate, is the assumption (9). As (9) is an inequality we have, from an experimental point of view, no possibilities of making a quantitative estimate of the error introduced by the use of postulate 2. However, we have formulated two postulates in order to make possible a formal discussion of certain problems and they are in this respect exactly valid, without any approximations. Thus we have the fundamental properties of $S$ independent of the experimental procedure, and so the logical background for the following discussions. But, as pointed out in the introduction, the form of an approach is to some extent a matter of taste. Thus, for instance, it may be argued that the state of $u$ is too vaguely introduced and that it would have been desirable to have this concept more rigorously defined. However, it is always necessary to begin somewhere using concepts that cannot be defined in any other manner than by referring to intuition: in Euclidian geometry the concepts "point" and "straight line" are introduced in this way. The same is true for the "particle" in classical mechanics, and, in the present theory, the two concepts "complete description of state" and "transition probability" are used in an analogous manner.

## 4. The Finite Tracer System

The approach of the preceding section may seem somewhat abstract and we shall try in this section to simplify the picture. This is done not only for visual purposes but also in order to simplify the mathematical treatment. Thus, we shall here concern ourselves with a special type of

tracer system, namely that in which the set $S$ in $Q$ contains a finite number of points only. We shall use the symbol (s) for such a system, where the number of states of $u \epsilon \mathbf{B}_{(s)}$ is finite. We may think of the general $S$ as divided into a finite but possibly large number of small "volume elements" $(ss)_i$, $i = 1, 2, \ldots, r$, so called *subsystems*, within which the $\bar{q}$-points are regarded as corresponding to one and the same state of $u$. This is certainly an approximation and we shall accept it in the same manner as one often divides, in elementary calculus, the integral into a finite number of small rectangles. Thus, to a given $S$ it seems reasonable to assume that we can always find an (s) the behavior of which, to any desired degree of accuracy, agrees with that of $S$. However, (s) itself is of course defined without approximation, and in that respect our treatment, as well as the results obtained, are non-approximate.

In the following we shall always assume (s) to be stationary (see page 127). Then, if we regard any $u \epsilon \mathbf{B}_{(s)}$ in the state $(ss)_i$, the probability of a transition to $(ss)_j$ can, according to postulate 2, be written as $P_{ij}(\Delta t)$, which thus is a function of $i$, $j$ and $\Delta t$ only. As $P_{ji}(\Delta t)$ is independent of $t$ we need not bother about the distribution or the total amount of the mother substance in (s); either these variables are constant or they do not have any influence upon the transition probabilities. As $u$ cannot be in two different states at the same time, the subsystems are in this respect mutually exclusive. Thus if, according to (4), $\mathbf{B}_i = \mathbf{B}_{(s)} \cdot (ss)_i$ = the set of $u \epsilon \mathbf{B}_{(s)}$ in $(ss)_i$ then, as in (3),

$$\mathbf{B}_i \cap \mathbf{B}_j \Big|_{i \neq} = \phi \tag{10}$$

Therefore, if we regard a particular $u$ in $(ss)_i$ we are in fact faced with a special kind of system, namely a so-called Markov chain (see, for instance, Feller, 1957, and Bartholomay, 1958a), which here is finite and stationary, i.e. it contains a finite number of mutually exclusive states and the transition probabilities are time-independent.

Let us regard an individual $u' \epsilon \mathbf{B}_{(s)}$ and let $\mathrm{P}'_i(t)$ = the probability that $u'$ is found in $(ss)_i$ at time $t$. Then, as all $u \epsilon \mathbf{B}_{(s)}$ have the same transition probability $P_{ji}(\Delta t)$ for the transition from $\bar{q}_i$ to $\bar{q}_j$ during the time interval $\Delta t$, we have

$$\mathrm{P}'_i(t + \Delta t) = \sum_{j=1}^{r} \mathrm{P}'_j(t) P_{ij}(\Delta t) \tag{11}$$

and

$$\lim_{\Delta t \to 0} P_{ij}(\Delta t) = \delta_{ij} \tag{12}$$

where

$$\delta_{ij} = \begin{cases} 0 \ (i \neq j) \\ 1 \ (i = j) \end{cases}$$

We can then write according to (11)†

$$\dot{P}'_i(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \left\{ P'_i(t + \Delta t) - P'_i(t) \right\}$$

$$= \lim_{\Delta t \to 0} \left\{ \sum_{\substack{j=1 \\ j \neq i}}^{r} P'_j(t) \frac{P_{ij}(\Delta t)}{\Delta t} + P'_i(t) \frac{P_{ii}(\Delta t) - 1}{\Delta t} \right\}$$

and so from (12)

$$\dot{P}'_i(t) = \sum_{j=1}^{r} P'_j(t) \lambda_{ij} \tag{13}$$

where $\lambda_{ij} = \dfrac{d}{d\Delta t} P_{ij}(\Delta t) \Big|_{\Delta t = 0}$, which means that

$$P_{ij}(\Delta t) = \delta_{ij} + \lambda_{ij} \cdot \Delta t + o(\Delta t) \tag{14}$$

The sum $\sum_{i=1}^{r} P_{ij}(\Delta t)$ is equal to the probability that $u'$ is in any (ss) whatever at time $t + \Delta t$, given that $u'$ is in $(ss)_j$ at time $t$. If the set of all $(ss)_i \subset (s)$ represents all the possible states in which $u'$ can be found, then $\sum_{i=1}^{r} P_{ij}(\Delta t) \equiv 1$ for $j = 1, 2, \ldots, r$, and so $u'$ can never leave (s). On the other hand if $\sum_{i=1}^{r} P_{ij}(\Delta t) < 1$ for some $j$-values, then there are states not contained in (s) where $u'$ may be found after time $\Delta t$, and so we have

$$\sum_{i=1}^{r} P_{ij}(\Delta t) \begin{cases} \equiv 1 \text{ when (s) is closed} \\ \leqslant 1 \text{ when (s) is not closed} \end{cases} \tag{15}$$

and as (15) is valid for all $\Delta t$, (14) gives

$$\lambda_{jj} \leqslant - \sum_{\substack{i=1 \\ i \neq j}}^{r} \lambda_{ij}, \ j = 1, 2, \ldots, r \tag{16}$$

where we have equality for all values of $j$ if and only if (s) is closed. (16) is an important relation to which we shall return in a following paper.

In order to obtain the physical meaning of our treatment we shall now concentrate our attention on (14). As follows from (14) the term

† In the following a dot over a symbol is used to indicate differentiation with respect to $t$.

$\lambda_{ij}\delta t \Big|_{i \neq j}$, where $\delta t$ stands for a very short time interval, can be interpreted as the probability that a $u \epsilon \mathbf{B}_j$ during time $\delta t$ shall move to $(\mathrm{ss})_i$. Thus $\lambda_{ij}\delta t \Big|_{i \neq j}$ represents the probability of a "success" in a conceptual sequence of Bernoulli trials repeated $(\mathrm{n}_j(t) + \mathrm{n}_i^0(t))$ times during the time interval $[t,\ t + \delta t]$ and where, according to (6),

$$\mathrm{n}_j(t) + \mathrm{n}_j^0(t) = \boldsymbol{N}\,(b_j(t) + b_j^0(t)) = \boldsymbol{N}\int_{(\mathrm{ss})_j} \{\sigma(\bar{q},\, t) + \sigma^0(\bar{q},\, t)\}\, d\bar{q} \qquad (17)$$

($\boldsymbol{N}$ = Avogadro's number).

This means that for the *expected number* of $u \epsilon \mathbf{B}_j$, which during $\delta t$ move from $(\mathrm{ss})_j$ to $(\mathrm{ss})_i$, we get $(n_j(t) + \mathrm{n}_j^0(t))\lambda_{ij}\,\delta t$. So if we write $\zeta_{ij}$ and $\zeta_{ij}^0$ for the macroscopic flux $(\mathrm{ss})_j \to (\mathrm{ss})_i$ of tracer and mother substance respectively, i.e. the number of gram atoms of $u \epsilon B_j$ and $u \epsilon B_j^0$ which in unit time move from $(\mathrm{ss})_j$ to $(\mathrm{ss})_i$, then

$$\zeta_{ij}(t) = \lambda_{ij}\, b_j(t),\ \zeta_{ij}^0(t) = \lambda_{ij}\, b_j^0(t) \qquad (18a)$$

and in particular for the tracer

$$\zeta_{ij}(t) = \lambda_{ij} b_j(t) = \zeta_{ij}^0(t) A_j(t) \qquad (18b)$$

$(A_j(t) = b_j(t)/b_j^0(t))$

This is the form of postulate 2 which is most common in the literature (see, for instance, Sheppard & Householder, 1951; Berman & Schoenfeld, 1956; and Robertson, 1957). $A_j$ is often known as the *specific activity* in $(\mathrm{ss})_j$.

There are several interesting features of the approach used here. For instance, the discussion of the "reaction order" has been avoided (see below). We understand that if the amount of tracer is sufficiently small (see (9)) then (18) is valid independently of the mechanism of the flux. It is also important to observe that $A_j$ is equal to $b_j/b_j^0$ and not equal to $b_j/(b_j^0 + b_j)$. Thus Sheppard & Householder (1951) assume as a principle that

$$\zeta_{ij} = (\zeta_{ij} + \zeta_{ij}^0)\,\frac{b_j}{b_j^0 + b_j} \qquad (19)$$

From an intuitive point of view (19) seems evident and it implies (18b). However, if instead of using postulate 2 we had begun with (19) then $\zeta_{ij}^0$ would have been equal to the flux of mother substance when the amount of tracer in $(\mathrm{ss})_j$ is equal to $b_j$, while from postulate 2 we have that $\zeta_{ij}^0$ is independent of $b_j$, and thus equal to the flux of mother substance

also when there is no tracer present in (s). But, to arrive at these conclusions from postulate 2 it has been necessary to assume (s) stationary. This is equivalent to the assumption that $\zeta^0_{ij}/b^0_j$ is constant, an assumption which seems to be evident only after the interpretation of $\zeta^0_{ij}/b^0_j$ as a probabilistic quantity. On the other hand, as this is a quite natural interpretation, we understand that on using (19) instead of postulate 2 we would probably have arrived at the same results. But somewhere on the way a basic assumption or postulate is necessary, and it seems then more satisfactory to begin with the postulate than to formulate it second-hand so to speak.

Also, the pronounced stochastic character of the approach is of some special interest. As (13) is valid for all $i$-values and all $u \epsilon B_{(s)}$ we have in matrix notation

$$\dot{\bar{P}}(t) = \boldsymbol{\lambda} \cdot \bar{P}(t) \tag{20}$$

where $\bar{P}(t)$ is a column vector $\begin{pmatrix} P_1(t) \\ \vdots \\ P_r(t) \end{pmatrix}$ and $\boldsymbol{\lambda}$ is a square matrix of order $r \times r$

$$\begin{pmatrix} \lambda_{11} & \lambda_{12} \ldots \lambda_{1r} \\ \lambda_{21} & \lambda_{22} \ldots \lambda_{2r} \\ \cdot & \cdot \quad \cdot \quad \cdot \quad \cdot \\ \lambda_{r1} & \lambda_{r2} \ldots \lambda_{rr} \end{pmatrix}$$

Let us now regard the "population" of $u \epsilon B_{(s)}$, i.e. the population of tracer atoms in (s). From (20) it then follows that if we know $\bar{P}(t)$ for $t = 0$, say, then $\bar{P}(t)$ is known for any value of $t \geqslant 0$ (see the assumption for $t$ in connection with (9) ). Thus we have a distribution of probabilities, where each probability $P_i(t)$ tells us the probability for any $u \epsilon B_{(s)}$ at $t = 0$ to be in $(ss)_i$ at time $t$. To each $u \epsilon B_{(s)}$ and each $(ss)_i$ there corresponds a binomial distribution with $P_i(t)$ as the probability of a "success", i.e. $u \epsilon B_i$ at time $t$, and $(1 - P_i(t))$ as the probability of a "failure", i.e. $u \epsilon B_i$ at time $t$. According to postulate 2 the probability for a transition of a particular $u \epsilon B_{(s)}$ is independent of what the other $u$'s are doing, and therefore all $u \epsilon B_{(s)}$ are statistically independent of each other. On considering the binominal distribution above, this then means that each $u \epsilon B_{(s)}$ can be regarded as an independent "trial", analogous to the discussion of the fluxes (page 130). Hence the expected number of "successes", i.e. the expected number of $u \epsilon B_i$, is given according to (17) by

$$N b_i(t) = N b_{(s)}(0) \cdot P_i(t) \tag{21}$$

$$\left(b_{(s)}(0) = \sum_{i=1}^{r} b_i(0)\right)$$

And so (20) implies†

$$\dot{\bar{b}}(t) = \lambda \cdot \bar{b}(t) \tag{22}$$

where $\bar{b}(t) = \begin{pmatrix} b_1(t) \\ \vdots \\ b_r(t) \end{pmatrix}$ is the column matrix of the expected number of $u \epsilon B_{(s)}$ in respective (ss) expressed in gram-atoms.

The resemblance between (20) and (22) is only formal: $\bar{P}(t)$ is a pure deterministic variable while $\bar{b}(t)$ is a stochastic one. Thus, there are no fluctuations associated with $P_i(t)$ while $b_i(t)$ is a mean value with corresponding standard deviation

$$D_i(t) = \sqrt{N\, b_{(s)}(0) P_i(t)(1 - P_i(t))}$$

which means that the standard deviation relative to the mean value is

$$\frac{D_i(t)}{N\, b_i(t)} = \frac{\sqrt{b_{(s)}(0)\,[P_i(t)(1 - P_i(t))]}}{b_i(t)\sqrt{N}} < \frac{\sqrt{b_{(s)}(0)}}{b_i(t)\sqrt{N}} \tag{23}$$

This is a kind of a general $\sqrt{N}$-rule of nature, which implies that on a macroscopic level such fluctuations can usually be neglected (see for instance, Bartholomay, 1958b, and Schrödinger, 1952).

Analogous arguments hold for the flux $\zeta_{ij}(t)$ too, when $\zeta_{ij}(t)$ is regarded as a conditioned stochastic variable, i.e. given that the content of $B_j$ at time $t$ is equal to $b_j(t)$ then the mean value of the flux of tracer from $(ss)_j$ to $(ss)_i$ is equal to $\zeta_{ij}(t)$ as given by (18), and the relative standard deviation is of the order $1/\sqrt{N b_j(t)}$. Now $\lambda_{ij} \mid_{i \neq j} \geqslant 0$ and so (16) implies that the diagonal elements of $\lambda$ are all nonpositive and, as is evident from (14), $|\lambda_{ii} \cdot \delta t|$ = the probability of a $u \epsilon B_i$ to leave $(ss)_i$ during time $\delta t$. Thus, $\zeta_{ii}(t) \leqslant 0$ and $|\zeta_{ii}(t)|$ = the expected flux of tracer from $(ss)_i$. Hence if there is no flux from $e$ into $(ss)_i$,

$$\dot{b}_i(t) = \sum_{j=1}^{r} \zeta_{ij}(t) \tag{24}$$

Regarding all $\zeta_{ij}(t)$ as independent stochastic variables, we understand that the standard deviation of $\dot{b}_i(t)$ may be considerable, especially for large $r$-values, even if the standard deviation of each $\zeta_{ij}(t)$ is small compared to $\zeta_{ij}(t)$ itself. However, $\dot{b}_i(t)$ is a macroscopic variable usually determined by observations of $b_i(t)$ at different $t$-values separated by finite intervals. Therefore, if we cannot observe any fluctuations in $b_i(t)$

† Observe that (20) is valid only for $u \epsilon B_{(s)}$ at $t = 0$, and it does not take into account any flux from $e$ into (s); the same restriction is valid also for (22).

then this will not be possible for $b_i(t)$ either.† Thus, in light of the discussion above we shall accept (18) and (22) (and then also (24)) without considering their statistical nature. Hence $\lambda_{ij}$ is given a statistical significance but the corresponding relations (18) and (22) are looked upon as being deterministic or phenomenological ones. But this is in part an approximation, and, as indicated in the introduction, due to experimental improvements it may some day be necessary to leave this deterministic view and take into consideration the statistical nature of the system. At present, however, distinguishing the fluctuations implied by the experimental procedure from those which are implied by the observed system itself seems to be an unsolved problem.

The (ss)-concept has no necessarily evident physical significance and we have introduced it mainly for simplifying the formal treatment. But at the same time there are several simple examples where the physical meaning of an (ss) is quite obvious. For instance, the simple system shown by Fig. 1 can be regarded as an (s) consisting of two (ss) only, or,

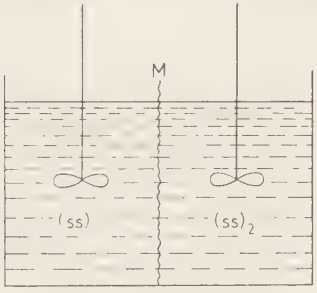



FIG. 1.

in common terminology, two *compartments with complete mixing*. Thus we have $(ss)_1$ and $(ss)_2$ containing water solutions of substances which can pass through the membrane M. With good stirring assumed, the probability that a certain particle in a given (ss) will hit M during time $\delta t$ can be regarded as independent of the geometrical position of the particle within that (ss). Even if this statement seems to contain some non-trivial problems we accept it as evident from an intuitive point of view. Thus the two compartments contain only one (ss) each, and we have from (18)

$$\mathcal{J}_{21}(t) \equiv \zeta_{21}(t) - \zeta_{12}(t) = \zeta^0_{21}(t)A_1(t) - \zeta^0_{12}(t)A_2(t)$$

or

$$\mathcal{J}_{21}(t) = \mathcal{J}^0_{21}(t)A_1(t) - \zeta^0_{12}(t)\Delta A_{21}(t) \tag{25}$$

$(\Delta A_{21}(t) \equiv A_2(t) - A_1(t))$

† In fact, on a microscopic level the time derivative of the content of $\mathbf{B}_i$ is equal either to zero or to infinity.

where $\mathcal{J}_{21}(t)$ and $\mathcal{J}^0_{21}(t)$ are the *net fluxes* of tracer and mother substance respectively from $(ss)_1$ to $(ss)_2$. This equation is essentially the same as that derived by Nims (1959, 2nd equation (36), page 333), which states that, unless $\Delta A_{21}(t) \equiv 0$, observation of the net flux of tracer gives no information about the net flux of the mother substance; one exception is of course the case where $\zeta^0_{12} \equiv 0$, $\zeta^0_{21} > 0$ or $\zeta^0_{12} > 0$, $\zeta^0_{21} \equiv 0$. However, the thermodynamic approach used by Nims† assigns no significance to the coefficient of $\Delta A_{21}(t)$. Whether the form of (25) is correct or not could probably be tested experimentally, but hitherto remarkably little basic experimental work in tracer dynamics has been published.

Finally, one more thing requires brief discussion, namely the physical significance of (9). Let us take, as a simple but illustrative example, the case when the flux $(ss)_j \to (ss)_i$ can be regarded as a reaction of order $n > 0$, i.e.

$$\zeta^*_{ij}(t) \equiv \zeta_{ij}(t) + \zeta^0_{ij}(t) = k_{ij}(b^*_j(t))^n$$

where $b^*_j(t) = b_j(t) + b^0_j(t)$. This equation should be regarded as a phenomenological one, and therefore no microscopic interpretation is required.

If from the beginning $b_j = 0$, then, for fixed $t$,

$$\zeta^*_{ij} = \zeta^0_{ij} \mid_{b_j = 0} = k_{ij}(b^0_j)^n$$

and thus

$$\delta\, {}^*_{ij} = nk_{ij}(b^0_j)^{n-1}\, \delta b^*_j + o(\delta b^*_j)$$

Let now a small amount $b_j$ of tracer be added to $(ss)_j$ such that $b_j \ll b^0_j$, hence

$$\delta\zeta^*_{ij} = n\, k_{ij}(b^0_j)^{n-1}\, b_j$$

The total flux of $u \epsilon \mathbf{B}_j$ is now

$$\begin{aligned}\zeta^*_{ij} &= \zeta^0_{ij}\,|_{b_j = 0} + \delta\zeta^*_{ij} = \zeta^0_{ij} + \zeta_{ij} \\ &= \zeta^0_{ij}\,|_{b_j = 0} + \delta\zeta^0_{ij} + \delta\zeta_{ij}\end{aligned}$$

where we evidently have

$$\zeta_{ij} = \delta\zeta_{ij} = \frac{1}{n}\,\delta\zeta^*_{ij} = k_{ij}(b^0_j)^{n-1}\, b_j$$

and

$$\delta\zeta^0_{ij} = \frac{n-1}{n}\,\delta\zeta^*_{ij} = (n-1)\, k_{ij}(b^0_j)^{n-1}\, b$$

† Nims works from the phenomenological equations of irreversible thermodynamics (see for instance Denbigh, 1951), and he assumes certain forms of the thermodynamic forces and also some special divisions of the net fluxes.

But, as $b_j \ll b_j^0$,

$$\zeta_{ij}^0 = \zeta_{ij}^0 \mid_{b_j=0} + \delta\zeta_{ij}^0 \approx \zeta_{ij}^0 \mid_{b_j=0} = k_{ij}(b_j^0)^{n-1} b_j^0$$

This is straightforward and on comparing with (18) we get

$$\lambda_{ij} = k_{ij}(b_j^0)^{n-1}, \; n > 0$$

However, in biological systems zero order reactions are not unlikely to occur, e.g. in enzymatic reactions where, according to the model of Michaelis and Menten, all the enzyme molecules are bound to the substrate. In such a case we have

$$\zeta_{ij}^{\bullet} = k_{ij}$$

or

$$\zeta_{ij} + \zeta_{ij}^0 = k_{ij}$$

Now $k_{ij}$ is proportional to the number of $u\epsilon\mathbf{B}_j$ in the substrate-enzyme complex, thus the probability that any $u\epsilon\mathbf{B}_j$ will be in this state is proportional to $k_{ij}'(b_j + b_j^0)$ where thus a $u\epsilon$(substrate-enzyme complex) is considered as belonging to $\mathbf{B}_j$. Repeating the arguments used in connection with (18) we get

$$\zeta_{ij}^0 = \frac{k_{ij}\, b_j^0}{b_j + b_j^0} \approx \frac{k_{ij}}{b_j^0}\, b_j^0, \quad \zeta_{ij} = \frac{k_{ij}\, b_j}{b_j + b_j^0} \approx \frac{k_{ij}}{b_j^0}\, b_j$$

as $b_j \ll b_j^0$. Instead of speaking of probabilities we could have used (19), which would immediately give the same results. Whichever way we choose, we have as an elementary conclusion:

*Theorem* 1. If the flux of $u\epsilon\mathbf{B}_{(s)}$ from $(ss)_j$ to $(ss)_i$ can be regarded as a reaction of order $n \geqslant 0$, then (18) is valid and

$$\lambda_{ij} = k_{ij}(b_j^0)^{n-1} \tag{26}$$

where $k_{ij}$ is the ordinary phenomenological rate constant.

Evidently, (26) is a more or less trivial consequence of (18), and, from the point of view of reaction kinetics, merely another way of writing postulate 2. However, there seems to be some confusion concerning the influence of the order of reaction on the tracer dynamic equations, not in the literature but at the laboratory. Further, the treatment above gives a good illustration of the approximative character of (18) and the necessity of (9): (9) is necessary except when $n = 1$, in case of which (18) is exactly valid. But, also if (9) is not necessary when $n = 1$ we must remember that $k_{ij}$ often is sensitive to variables such as ionic strength and pH, variables that may be influenced by the addition of tracer. Thus, even when $n = 1$ the amount of tracer must be kept as small as possible. The necessity of this condition is also evident from another point of view: at high concentrations there will be an interaction between the tracer

particles. The importance of the mother substance may be, among others, that to a certain extent it prevents such interactions.

## 5. The Concept of Turnover

Let $C_\nu$ be a compartment as defined on page 125, and let $d\lambda(\bar{q}', \bar{q})\delta t$ be equal to the probability of a transition during time $\delta t$ from $\bar{q}$ to $\bar{q}'$.† If $Q'_\nu = Q - C_\nu$, i.e. the complement of $C_\nu$ relative the $Q$-space, we define for $\bar{q}\epsilon C_\nu$

$$\lambda(\bar{q}, \nu) \equiv \int_{Q'_\nu} d'\lambda(\bar{q}', \bar{q}) \tag{27a}$$

where the integration is performed with respect to $\bar{q}'$. Since all the transitions $\bar{q} \to \bar{q}'$ are mutually exclusive for different $\bar{q}'$, $\lambda(\bar{q}, \nu)\delta t$ is equal to the probability that a $u$ in the state $\bar{q}\epsilon C_\nu$ during time $\delta t$ shall move to one of the states belonging to $Q'_\nu$. We call $\lambda(\bar{q}, \nu)$ the *turnover coefficient* of $\mathbf{B}_\nu = \mathbf{B}_S \cdot C_\nu$ (see page 126) and define the *tracer turnover function* of $\mathbf{B}_\nu$ as

$$\varphi_\nu(t) \equiv \int_{C_\nu} \lambda(\bar{q}, \nu) f_\nu(\bar{q}, t)\, d\bar{q} \tag{28}$$

where $f_\nu(\bar{q}, t)$ is the relative content density as defined by (7). The significance of $\varphi_\nu(t)$ is evident: $\varphi_\nu(t)\delta t$ is the expected probability at time $t$ of a $u\epsilon B_\nu$ to leave $C_\nu$ during a short time interval $\delta t$. It should be remembered that we assume $u$ given with a fixed atomic number, and that in this respect all our treatment is hitherto and throughout concerned with one and the same $u$. Thus, if we are dealing with $u$'s having different atomic numbers, then it would be necessary to indicate in the symbols which atomic number they refer to.

Now, to each $C_\nu$ there exists one and only one $Q'_\nu$ and one and only one

$$C'_\nu = S - C_\nu$$

and we define the *turnover coefficient of* $\mathbf{B}'_\nu = \mathbf{B}_S \cdot C'_\nu$ as

$$\lambda(\bar{q}', \nu) \equiv \int_{C'_\nu} d\lambda(\bar{q}, \bar{q}') \tag{27b}$$

and in terms of that the *tracer turnover function of* $\mathbf{B}'_\nu$

$$\varphi'_\nu(t) \equiv \int_{C'_\nu} \lambda(\bar{q}', \nu) f_\nu(\bar{q}', t)\, d\bar{q}' \tag{29}$$

† This means that we assume: if $P(\bar{q}', \bar{q}, \Delta t, t)$ is the transition probability in postulate 2 corresponding to the two points $\bar{q}$ and $\bar{q}'$ then $\partial^n P(\bar{q}', \bar{q}, \Delta t, t)/(\partial \Delta t)^n \Big|_{\substack{t \geq 0 \\ \Delta t = 0}}$ exists for $n = 0, 1, 2, \ldots$ and for any $\bar{q}$ and $\bar{q}'$ belonging to $Q$.

Here $\varphi'_\nu(t)\delta t$ = the expected probability at time $t$ of a $u \epsilon B_\nu$ to move into $C_\nu$ during a short time interval $\delta t$.

We thus regard $Q$ as divided into the two sets of points $S$ and $e$, after which we have selected an arbitrary compartment $C_\nu$ of $S$, i.e. $C_\nu \subset S$. Corresponding to this $C_\nu$ there are two complements: $Q'_\nu$ with respect to $Q$ and $C'_\nu$ with respect to $S$. Further, in terms of the coefficients of turnover $\lambda(\bar{q}, \nu)$ and $\lambda(\bar{q}', \nu)$ we define the tracer turnover functions $\varphi_\nu(t)$ and $\varphi'_\nu(t)$ such that, when $b_\nu(t)$ and $b'_\nu(t)$ are the contents of $B_\nu$ and $B'_\nu$ respectively at time $t$, the expected fluxes of tracer from $C_\nu$ and from $C'_\nu$ into $C_\nu$ are given by $\varphi_\nu(t)b_\nu(t)$ and $\varphi'_\nu(t)b'_\nu(t)$ respectively. Hence, if $\zeta_{\nu e}(t)$ = the flux of tracer from $e$ to $C_\nu$, we have the ordinary continuity equation of the form

$$\dot{b}_\nu(t) = \varphi'_\nu(t)b'_\nu(t) - \varphi_\nu(t)b_\nu(t) + \zeta_{\nu e}(t) \tag{30}$$

where as usual we regard the fluxes not as stochastic but as deterministic variables.

Of interest is the tracer turnover function $\varphi_\nu(t)$. Now, according to postulate 2 the tracer and the mother substance have the same turnover coefficient $\lambda(\bar{q}, \nu)$, hence the *turnover function of* $\mathbf{B}_\nu$ is

$$\varphi^0_\nu(t) \equiv \int_{C_\nu} \lambda(\bar{q}, \nu) f^0_\nu(\bar{q}, t)\, d\bar{q} \tag{31}$$

Let us make the following definition: $S$ is said to be *perfect* if it is not only stationary (which here implies that $\dot{\lambda}(\bar{q}, \nu) \Big|_{\bar{q} \epsilon C_\nu \subset S} \equiv \dot{\lambda}(\bar{q}', \nu) \Big|_{\substack{\bar{q}' \epsilon C_\nu' \\ C_\nu \subset S}} \equiv 0$) but also if $\dot{o}^0(\bar{q}, t) \Big|_{\substack{t \geqslant 0 \\ \bar{q} \epsilon S}} \equiv 0$. Thus, if $S$ is perfect then $f^0_\nu(\bar{q}, t) \Big|_{\bar{q} \epsilon C_\nu \subset S} = \pi^0_\nu(\bar{q}) \Big|_{\bar{q} \epsilon C_\nu \subset S}$ and we get from (31)

$$\varphi^0_\nu(t) \equiv \int_{C_\nu} \lambda(\bar{q}, \nu)\pi^0_\nu(\bar{q})\, d\bar{q} = \lambda^0_\nu \tag{32}$$

Hence, when $S$ is perfect $\varphi^0_\nu(t) \Big|_{t \geqslant 0} = \lambda^0_\nu$ and we shall call this parameter the *turnover factor of* $\mathbf{B}_\nu$, which thus refers to the mother substance of $\mathbf{B}_\nu$. This makes it possible for us to define the *turnover of* $\mathbf{B}_\nu$ as

$$h^0_\nu = \lambda^0_\nu \cdot b^0_\nu \tag{33}$$

which is the flux of mother substance from $C_\nu$.

When $S$ is perfect the fluxes of the mother substance are all constant with respect to $t$, and only then can we speak of a turnover. However

common the concept of turnover may be, it is very seldom rigorously defined and in the literature different opinions concerning its meaning can be found. For instance, Bray & White (1957) define the turnover in terms of "the amount of substance removed in unit time", i.e. the flux out of $C_\nu$, while Robertson (1957) seems to be of the opinion that one should use two turnovers: the flux into $C_\nu$ and the flux out, in case they are different. Now, if the turnover is intended to describe certain properties of $B^0_\nu$, then it is not justified to regard the influx as the turnover: the influx is equal to $(\varphi^{0\prime}_\nu(t)b^{0\prime}_\nu(t) + \zeta^0_{\nu e}(t))$ and thus dependent on variables or parameters that are characteristics of $B^{0\prime}_\nu$ and $\zeta^0_{\nu e}(t)$. But as we are discussing here a perfect $S$, this problem is more or less of academic interest only, as in such an $S$ the influx from $Q'_\nu$ must be equal to $h^0_\nu$ defined by (33). The assumption that $S$ is perfect may perhaps at a first glance look quite restrictive. However, we have assumed that $S$ is stationary, otherwise no description of the dynamic state of $S$ in terms of constant parameters would be possible (see page 127). Further, from theorem 1 we have that, unless we assume all the fluxes to follow first order kinetics, variations in $b^0_\nu$ are likely to influence the turnover coefficient $\lambda(\bar{q}, \nu)$, hence, if we actually want to speak of a turnover factor, $b^0_\nu$ must in general be constant. Also necessary for the existence of $\lambda^0_\nu$ is that $f^0_\nu(\bar{q}, t) = \pi^0_\nu(\bar{q})$ for all $\bar{q}\epsilon C_\nu$, and we therefore understand that the assumption that $C_\nu$ is perfect is, if not quite necessary, difficult to avoid. As this implies that the influx and efflux of the mother substance in $C_\nu$ must balance each other, a moment's consideration will make evident that also $C'_\nu$ must be perfect, otherwise theoretically possible but physically unrealistic assumptions concerning the properties of $B^{0\prime}_\nu$ and $\zeta^0_{\nu e}(t)$ are necessary. However, there seem to exist systems where a perfect state is never attained, e.g. some endocrine systems (Danziger & Elmergreen, 1957), and such systems are not, of course, included in our treatment.

We have thus defined a turnover or, perhaps the more rational concept, a turnover factor. The definition is rigorous, and applicable to a large class of systems: there are no restrictions on the choice of the compartment $C_\nu$, e.g. we may take $C_\nu$ to be the non-protein alanine of the liver, or the total liver alanine, or the whole set of amino acids in the liver and so on. Now, if we have chosen $C_\nu$ to be the non-protein alanine in the liver and we regard the case where $u$ = nitrogen, then $B^0_\nu$ is equal to the set of all naturally occurring nitrogen atoms in this compartment. It is then evident that the corresponding $h^0_\nu$ may be quite different from the one, for instance, that corresponds to $u$ = oxygen. Thus, when we concern ourselves with a certain metabolite, the molecule of which contains atoms with different atomic numbers, it is in general not possible to speak of only one turnover of that metabolite: the turnover must be specified with respect to a special

atom or a kind of atoms contained in the molecule of the metabolite considered.

It is important that the turnover concept introduced here refers to an atom chosen at random from the whole population of the atoms under consideration; thus it can be used rigorously without referring to any detailed model such as a "precursor map". A simple but important fact, which also makes the significance of the (ss) concept more evident, is that according to (8)

$$\varphi_i(t) \equiv \int_{(ss)_i} \lambda(\bar{q}, t) f_i(\bar{q}, t)\, d\bar{q} = -\lambda_{ii} \int_{(ss)_i} f_i(\bar{q}, t)\, d\bar{q} = -\lambda_{ii}$$

$$= -\lambda_{ii} \int_{(ss)_i} f_i^0(\bar{q}, t)\, d\bar{q} \equiv \varphi_i^0(t) \tag{34}$$

Hence, for any subsystem there exists always a turnover factor and a "tracer turnover factor" such that $\lambda_i^0 = \lambda_i = -\lambda_{ii}$, where $\lambda_{ii}$ has the same significance as in the preceding section.

If now $\lim_{t\to\infty} f_\nu(\bar{q}, t) = \pi_\nu(\bar{q})$ exists for all $\bar{q}\epsilon C_\nu$ and for any $C_\nu \subset S$, $S$ is said to be a *limiting* $S$. If $S$ is a limiting $S$ then

$$\lim_{t\to\infty} \varphi_\nu(t) = \lambda_\nu, \quad \lim_{t\to\infty} \varphi_\nu'(t) = \lambda_\nu' \tag{35}$$

both exist for any $C_\nu \subset S$, and we shall refer to $\lambda_\nu$ as the *tracer turnover factor of* $\mathbf{B}_\nu$. On comparing (28) with (31) we immediately have:

*Theorem* 2. If $S$ is a perfect and limiting $S$, the turnover factor $\lambda_\nu^0$ is equal to the tracer turnover factor $\lambda_\nu$ when (1) $\pi_\nu(\bar{q}) \mid_{\bar{q}\epsilon C_\nu} = \pi_\nu^0(\bar{q}) \mid_{\bar{q}\epsilon C_\nu}$, and/or

(2) $\dfrac{d}{d\bar{q}} \lambda(\bar{q}, \nu) \mid_{\bar{q}\epsilon C_\nu} = 0.$

Condition (2) is an analogue to the "lumping effect" first described by Sheppard & Householder (1951) and later in a somewhat more general manner by Hart (1955) and which in light of (34) can be expressed thus: when $\lambda(\bar{q}, \nu)$ is constant for all $\bar{q}\epsilon C_\nu$, then $S$ behaves as if $C_\nu$ is a subsystem. However, this theorem faces us with the question of how to perform a tracer experiment in such a way that the tracer "represents" the mother substance. Together with the problem of the determination of $\lambda_\nu$ from experimental data, this question forms the foundation of the fundamental problem of tracer kinetics and to this we shall return in a later communication.

I wish to express my thanks to Fil. dr. Ulf Uhlhorn, Department of Theoretical Physics, Stockholm, for stimulating discussions and criticism on various points of this investigation, and also to Dr. Anthony F. Bartholomay, Harvard Medical School, Boston, for critical discussions concerning an early stage of this work.

## REFERENCES


Bartholomay, A. F. (1958a). *Bull. math. Biophys.* **20,** 97.
Bartholomay, A. F. (1958b). *Bull. math. Biophys.* **20,** 175.
Bartholomay, A. F. (1959). *Bull. math. Biophys.* **21,** 363.
Bergner, P.-E. E. (1959). *Exp. Cell. Res.* **17,** 328.
Bergner, P.-E. E. (1960). *Exp. Cell. Res.* **20,** 579.
Berman, M. & Schoenfeld, R. (1956). *J. appl. Phys.* **27,** 1361.
Berman, M. & Schoenfeld, R. (1960). *Exp. Cell. Res.* **20,** 574.
Bigeleisen, J. (1949). *Science,* **110,** 14.
Bray, H. G. & White, K. (1957). "Kinetics and Thermodynamics in Biochemistry". J. & A. Churchill Ltd., London.
Danziger, L. & Elmergreen, G. (1957). *Bull. math. Biophys.* **19,** 9.
Denbigh, K. G. (1951). "The Thermodynamics of the Steady State". Methuen & Co. Ltd., London.
Feller, W. (1957). "An Introduction to Probability Theory", 2nd Edition. J. Wiley, New York.
Hart, H. E. (1955). *Bull. math. Biophys.* **17,** 87.
Hart, H. E. (1957). *Bull. math. Biophys.* **19,** 61.
Hart, H. E. (1958). *Bull. math. Biophys.* **20,** 281.
Hart, H. E. (1960). *Bull. math. Biophys.* **22,** 41.
Nims, L. F. (1959). *Yale J. Biol. Med.* **31,** 373.
Rescigno, A. (1954). *Biochim. biophys. Acta,* **15,** 340.
Rescigno, A. (1956). *Biochim. biophys. Acta,* **21,** 111.
Rescigno, A. (1960). *Biochim. biophys. Acta,* **37,** 463.
Robertson, J. S. (1957). *Physiol. Rev.* **37,** 133.
Schrödinger, E. (1952). "Statistical Thermodynamics". Cambridge University Press.
Sheppard, C. W. & Householder, A. S. (1951). *J. appl. Phys.* **22,** 510.
Solomon, A. K. (1949). *J. clin. Invest.* **28,** 1297.

*J. Theoret. Biol.* (1961) **2**, 141–179

# Human Colour Vision and the Perception of Blue

E. N. Willmer

*Physiological Laboratory, University of Cambridge, England*

(*Received* 20 *September* 1960)

1. The properties of the "blue mechanism" are described. It is shown that they differ in many ways from those of the "red and green mechanisms".
2. The possible contributions of rods to photopic vision are outlined and discussed.
3. There are strong arguments both for and against the view that rods are concerned with the "blue mechanism".
4. Evidence is put forward for the inhibition of rods by cones.
5. It is suggested that in photopic conditions the residual activity of the scotopic mechanism, i.e. the activity of the rods which escapes inhibition, constitutes the basis for the blue mechanism.
6. Hypothetical luminosity curves for the blue mechanism in protanopes, deuteranopes and in normal subjects have been plotted on this basis. It is deduced that the blue mechanism could be dependent upon rods being inhibited by cones in accordance with rules similar to those which Hartline has found to hold for the somewhat similar inhibitory processes which occur in the retina of *Limulus*. Possible explanations for the existence of two blue mechanisms (the $\pi_1$ and $\pi_3$ mechanisms) as revealed by Stiles are put forward. The hypothesis accounts reasonably well both qualitatively and quantitatively for many of the main features of colour vision.
7. Approximate luminosity curves for various retinal locations are predicted in terms of the estimated numbers of receptive fields for rods and cones in those areas, on the assumptions that luminosities are directly additive and that, probably because of the average areas of their receptive fields, about 7·5 rods are equivalent to one cone, in so far as luminosity is concerned. These curves are directly comparable with those actually found.
8. Various anomalous observational data appear to be explicable in terms of the hypothesis which is outlined.
9. The hypothesis is certainly not without its difficulties; there are at least two very strong arguments based on the accepted principles of colorimetry which will have to be satisfactorily answered before it can be accepted.
10. On the other hand, if rods play no part in colour vision three equally difficult questions must be answered. They are:—
(*a*) How is the response from rods eliminated from photopic vision?
(*b*) Why do the hypothetical "blue" cones have such different properties from the "red" and "green" cones, if they are themselves cones?

(*c*) Why is there no direct anatomical or physiological evidence for the existence of these "blue" cones, if they are so different in kind ?

11. The ideas put forward in this paper are intended more to stimulate thought on the nature of the blue mechanism, which still remains the outstanding problem in the physiology of colour vision, than to constitute an exact description of the mechanism

## Introduction

Human colour vision is now firmly believed to depend upon three separate receptor mechanisms. In normal subjects the three can be reduced to two by confining vision very strictly to the central fovea (i.e. by directing the gaze steadfastly at objects subtending less than 0·5° at the eye). By this means, the so-called blue mechanism can be effectively excluded (König, 1894; Willmer, 1944; Willmer & Wright, 1945). The eye is then dichromatic and tritanopic. In red-green-blind subjects the simplification can be carried one stage further by the same technique and single mechanisms can then be isolated (Willmer, 1949b). The red-blind subject, or protanope, shows only the green mechanism in his central fovea and at least some green-blind subjects, or deuteranopes, show only the red mechanism when their vision is similarly confined to the foveal centre. Such normally dichromatic subjects become monochromatic.

In 1955 the results of an analysis of central foveal vision were published (Willmer, 1955) and it was shown that much of the reduced normal vision characteristic of this area (i.e. the dichromacy and tritanopia, could be explained on the basis of these two mechanisms, the green (*P*) and the red (*D*), each of which was thought to depend upon a single carotenoid-protein type of pigment. It was shown that the sensations of colour (i.e. the saturation and hue) which could be evoked from this area were closely related to the differences between the ordinates of the spectral sensitivity curves for the two mechanisms (as determined by the data obtained from the foveal centre of colour-blind subjects) when these curves were plotted on a logarithmic scale and so linked as to have the same sensitivity at the observed neutral, colourless or white point in the spectrum, which was found to lie at 574 m$\mu$ (Fig. 1). Since that time, the work of Rushton (1957, 1958) on the pigments of the fovea of normal and colour-blind subjects has confirmed the belief that the *P* mechanism (i.e. the green mechanism, which is present in isolation in the protanope or red-blind) does indeed depend on a single pigment. His data for the *D* mechanism are, however, not so clear-cut, and it is not yet certain from them whether this mechanism depends entirely on a single pigment or not.

The spectral sensitivities of these two mechanisms, as isolated in the central fovea and represented by the *P* and *D* curves of Fig. 1, are closely

similar to those which Stiles (1953) has called the $\pi_4$ and $\pi_5$ mechanisms, on the basis of his work on increment thresholds with two-colour fields.

By this technique, the eye is adapted to a certain intensity of a coloured

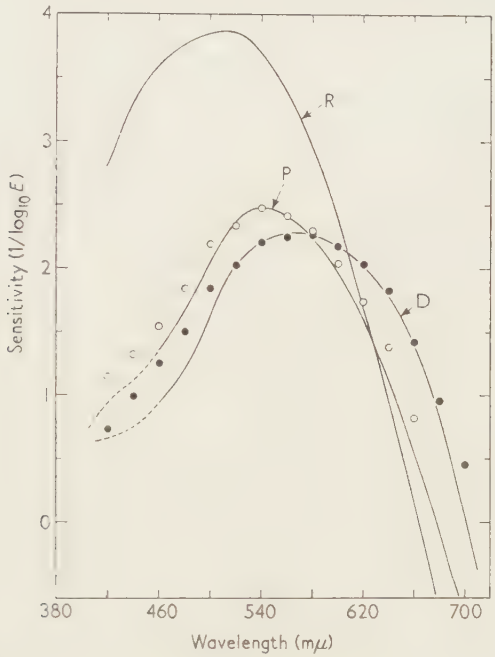



FIG. 1. The sensitivity curves of the three known receptors in the human eye.

*R*. The rod mechanisms (curve from Flamant & Stiles, 1948) dependent on rhodopsin.

*P*. The central foveal luminosity curve for the protanope (Willmer, 1955). The receptors are probably dependent on chlorolabe (Rushton, 1958b). The open circles represent the $\pi_4$ mechanism of Stiles (1953).

*D*. The central foveal luminosity for the deuteranope (Willmer, 1955). These receptors may depend on a mixture of chlorolabe and erythrolabe (Rushton, 1958b). The filled circles represent the $\pi_5$ mechanism of Stiles (1953).

The *P* and *D* curves have been linked at 574 mμ, which is the neutral point in the spectrum for the dichromatic central fovea; at this point the sensation changes from green to red (Willmer, 1955).

The *R* and *D* curves have been linked at 610 mμ, since the directional effect of light which only affects cones becomes detectable from about this point to the red end of the spectrum (Flamant & Stiles, 1948).

field and a test patch of another colour is superimposed upon it. By measuring the luminance of the test field necessary for the superimposed stimulus to be perceived, and by varying the intensity of the adapting field, it is possible to study the effects of the adapting wavelength on the sensitivity of the mechanisms used in the perception of the test patch, and thus by varying the wavelengths of the two fields, to determine the sensitivity of these mechanisms to the various parts of the spectrum. For

example, if vision of a given test colour were to depend upon two mechanisms, then with no adapting field a threshold could be found for the perception of the test colour. If then an adapting field were turned on this would probably depress the sensitivity of one of the mechanisms more than the other (depending on the wavelengths used) and so the threshold for the perception of the test object, would be raised, and, as the intensity of the adapting light was further increased, it would continue to be raised in a regular manner until the adapting field began to affect the sensitivity of the second mechanism. There would then be a discontinuity in the intensity-threshold curve. Thus it is possible to find that intensity of the adapting field at any wavelength which will just alter the sensitivity of the various mechanisms involved in vision at different wavelengths: Stiles has shown by this method that there are at least five such mechanisms ($\pi_1$ to $\pi_5$); i.e. there are five mechanisms that can have their thresholds altered by adaptation in the manner described. The technique does not indicate whether these mechanisms are actual receptors or some other susceptible levels in the neural pathways of vision.

In the earlier paper on central foveal vision it was suggested that, since the frequency of nerve impulses in the nerves leading from photoreceptors is, in general, much more closely related to the logarithm of the intensity of the stimulus than to the intensity directly, the difference between the ordinates of the sensitivity curves (plotted on a logarithmic scale), is really a measure of the difference between the frequencies of impulses set up in the two pathways, and this difference was shown to be related closely to the saturation of the colour. Thus, in vision by the foveal centre, when the $P$ mechanism initiates a higher impulse-frequency than the $D$ mechanism, sensations of blue-green are set up; when the $D$ mechanism is the more active, various degrees of redness result. The saturation or vividness of the colour, in each case, increases with the extent of the excess activity of the one mechanism over the other. When the two mechanisms are responding equally, a sensation of "no colour" is produced. The saturation of the colour thus depends upon the magnitude of the difference; and hue discrimination, in its simplest form, depends upon the rate of change of this magnitude with wavelength.

Though, at the time, this was a purely hypothetical concept based on human central foveal vision, the recent work of Svaetichin and his colleagues (MacNichol, Macpherson & Svaetichin, 1957; Svaetichin & MacNichol, 1958) on the retinae of fish has provided experimental evidence for the existence of a mechanism, in these animals at least, which could be interpreted to be working in exactly this way. They have shown, by inserting micro-electrodes into the retinae of certain fish, that there are cells or structures that show an increased negative potential in response to

illumination with light from the blue end of the spectrum and a positive potential with light from the red end of the spectrum. Clearly, if other cells were influenced by these increased positive or negative potentials, their responses would give precisely those required for such a mechanism as that postulated for human foveal colour-vision. It is of course true that this evidence, which is the only direct evidence, is derived from fish and not from man, but, at least, it is now clear that the suggested mechanism is certainly possible, at the neurological level, since it actually occurs in other animals.

Furthermore, MacNichol, Macpherson and Svaetichin have shown that among the horizontal cells, or closely related structures, there are some that, upon illumination, show slow and sustained negative potentials of an extent which in scotopic conditions varies spectrally in the same manner as the sensitivity of rhodopsin. Under photopic conditions, in some species at least, the maximal spectral sensitivity for these negative potentials shifts towards the red as in the typical Purkinje shift. The magnitude of these potentials is, at any one wavelength, rather closely related to the logarithm of the intensity of the light. They regard these cells as luminosity detectors, and it is interesting that they are distinct from those which could be interpreted as colour detectors.

In the paper on foveal vision (Willmer, 1955) it was also shown that, by making certain simple assumptions, the central foveal luminosity curve could be derived from the spectral sensitivity curves of the two mechanisms (the $P$ and $D$ mechanisms). It was assumed that the light of any wavelength affected the two mechanisms in such a way that each produced its own characteristic effect on the first synaptic connections leading to the bipolar cell concerned with luminosity. The effect produced depended partly on the sensitivity of the receptor to each particular wavelength and partly on another factor, which was probably the number of receptors of each type involved, i.e. on the number of receptors whose responses were transmitted to the particular bipolar cell. The effects produced by each receptor then added together in stimulating the bipolar cell; the response of the bipolar cell, when thus stimulated by the changes in the first synaptic layer, was proportional to the logarithm of the added effects. The data were such as to suggest that, in the central fovea, the $D$ receptors were either much more numerous than the $P$ receptors ($P : D = 2 : 3$) or that, for some other reason, they contributed much more to the common synaptic pool than the $P$ receptors. The luminosity curve was found to result from the expression:—

$$\log(xk_1p + yk_2d)$$

where $p$ and $d$ represent the actual intensities of the stimulus at any given

wavelength in so far as it affected the $P$ and $D$ receptors respectively: $k_1$ and $k_2$ are factors relating the stimulus to the response (they were actually made equal to 1 since it was assumed that this was likely to be very much the same for both types of cone) and $x$ and $y$ represent the relative numbers of the two types of receptor. Empirically it was found that a good fit with the experimentally determined luminosity curve was obtained when $x = 2$ and $y = 3$.

The experimentally observed differences in the shape of the luminosity curve (Thomson & Wright, 1947) with central fixation, and at 20′ and 40′ off-centre, were best fitted with values of $xk_1$ and $yk_2$ in the following ratios: central, 2 : 3; 20′ off-centre, 2 : 2; and 40′ off-centre 2 : 1 respectively. These effects could be explained either by differences in relative numbers ($x$ and $y$) or by variations in the effectiveness of the stimuli on each of the receptors ($k_1$ and $k_2$) or both; the first explanation seems the more probable. All these calculations were based on the central foveal measurements only, and they neglect the presence of the "blue" mechanism, though certain aspects of this mechanism are certainly beginning to be appreciable at 40′ off-centre.

In the following pages, vision in the whole eye, i.e. trichromatic vision, will be considered on the same lines as those adopted for vision by the central fovea and in relation to the spatial distribution of the known receptors in the human retina. The nature of the blue mechanism is of course the main enigma, and some of its properties and peculiarities will first be outlined.

### The Properties of the Blue Mechanism

1. From colour-mixture data (Wright, 1946), the blue mechanism has its maximum sensitivity at a wavelength about 450 m$\mu$, and contributes little or nothing to vision at wavelengths longer than about 540 m$\mu$ in the normal eye, about 540 m$\mu$ in deuteranopes, and about 550 m$\mu$ in protanopes.

2. From increment-threshold data (see p. 143) in the normal eye, Stiles (1953) has established the existence of two separate mechanisms, $\pi_1$ and $\pi_3$ (Figs. 2 and 3), both maximally sensitive in the neighbourhood of 450 m$\mu$. Of these, $\pi_3$ could perhaps be initially dependent directly upon a single pigment of the rhodopsin group, since its spectral sensitivity curve is of the right shape (Dartnall, 1957), but the spectral sensitivity of the $\pi_1$ mechanism would be difficult to reconcile with such an hypothesis on account of the curious hump between wavelengths 580 and 600 m$\mu$. If the visual pigment of this mechanism is, as in most other visual mechanisms, a carotenoid-protein, it must be "screened" in some way. Alternatively the "mechanism" is a "neural" path and not directly receptoral.

3. The blue mechanism can be eliminated from normal vision by using

only the central fovea. Whether this "elimination" depends on actual absence of the mechanism from this area, or on the action of some inhibitory mechanism which comes into play with central fixation, is unknown.

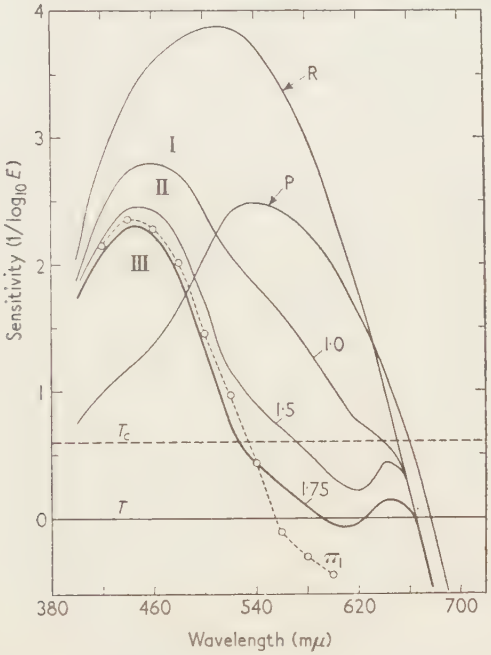



FIG. 2. Curves showing the effect of inhibition of the rods by the *P* cones, as in the protanope, in accordance with the type of inhibition which has been shown by Hartline and his colleagues (Ratliff, Miller & Hartline, 1958) to occur in the eye of *Limulus*.

*R*. Rod sensitivity curve; *P*. Sensitivity curve of *P* cones (see Fig. 1); *T*, threshold above which the mechanisms produce responses in proportion to the ordinates.

Curve I. $B = R - 1{\cdot}0\,(C - C_0)$
Curve II. $B = R - 1{\cdot}5\,(C - C_0)$
Curve III. $B = R - 1{\cdot}75\,(C - C_0)$

Where $B$ = ordinates expressing the sensitivity of the blue mechanism, $R$ = ordinates of the $R$ curve measured from $T$, $C - C_0$ = ordinates of the $P$ curve measured from $T_C$, the point at which the cone response becomes effective as an inhibitory stimulus.

o – – o, the $\pi_1$ mechanism of Stiles (1953).

There is some evidence that very small and stationary images falling elsewhere on the retina also tend to be seen in a somewhat tritanopic manner, i.e. without the intervention, or with much less intervention, of the blue mechanism (Hartridge, 1947). A completely stabilized image anywhere on the retina soon ceases to arouse any sensation of colour (Ditchburn & Ginsborg, 1952; Riggs, Ratliff, Cornsweet & Cornsweet, 1953; Ditchburn, 1958).

4. Apart from this elimination from the central fovea, the blue mech-

anism shares with that for yellow the widest distribution over the retina. The visual fields for blue and yellow are considerably larger than those for red and green, and, when bright enough lights are used, they probably extend over the whole retina. Indeed, the retinal periphery has, under photopic conditions, properties that are closely similar to those associated with dichromacy, and probably with deuteranopia.

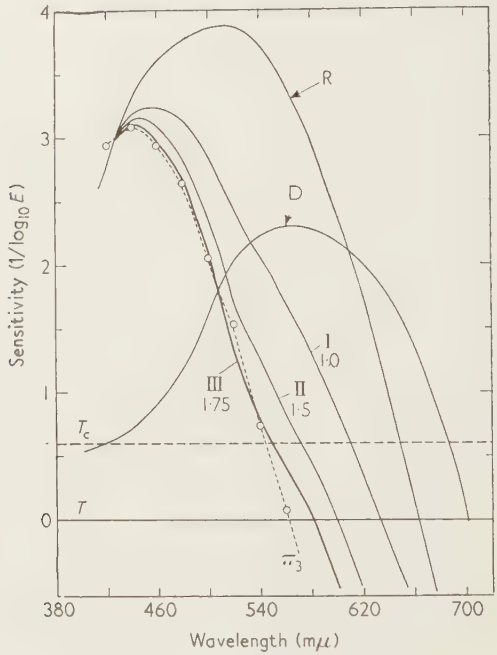



FIG. 3. Curves showing the effect of inhibition of the rods by the *D* cones of the deuteranope. The curves are derived in exactly the same manner as those in Fig. 2. o – – o, the $\pi_3$ mechanism of Stiles (1953).

5. In both deuteranopes (green-blind subjects) and protanopes (red-blind subjects) the blue mechanism is at least normally developed, and these subjects see all colours in shades of either blue (violet?) or yellow. The spectral colour-mixture curves for protanopes, deuteranopes and normal subjects indicate that the blue mechanism contributes proportionally to very different extents in each of these types; nevertheless the sensation of blueness is, in general and like that of yellowness, only very rarely absent or grossly abnormal. This suggests that the mechanisms responsible for these sensations depend upon some much more constant and stable features in the eye or visual pathway than do the red and green

mechanisms, which themselves seem often to be "fused" into a common mechanism, namely the yellow mechanism itself.

6. While deuteranopia and protanopia, in which the abnormalities certainly involve the cones or their connections, are inherited as sex-linked allelomorphs, tritanopia (blue-blindness) is inherited quite separately as an autosomal recessive (Kalmus, 1955). The gene for the development of the structures responsible for the blue mechanism is thus carried in a different chromosome from the genes for the red-green complex.

7. Adaptation to strong yellow light (of wavelength about 575 m$\mu$) produces very little change in the appreciation of hue when vision is confined to the central fovea, but, with larger fields, a state of "violet monochromatism" is produced (Brindley, 1953). In this state, all wavelengths between 400 m$\mu$ and 500 m$\mu$ appear identical and as a very saturated violet. In this condition the blue mechanism is probably acting in

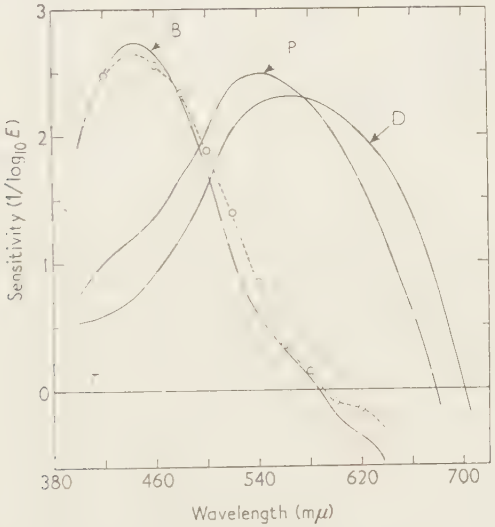



FIG. 4. The proposed sensitivity curves for the three photopic mechanisms of the normal eye.

$P$ = $P$ cones; $D$ = $D$ cones (see Fig. 1).

$B$ = the mean curve of the two curves marked III in Figs. 2 and 3. This represents the probable residual activity of the rods after inhibition by a mixed population of $P$ and $D$ cones.

o – – o, $Log_{10}$ increment threshold curve after adaptation to a high intensity of yellow (578 m$\mu$) (Stiles, 1953) which isolates the blue mechanism (Brindley, 1953).

great excess over the other mechanisms but, unfortunately, its spectral sensitivity cannot easily be measured because, for some unexplained reason, the judgement of luminosity is extremely bad under these conditions. This quality of the blue mechanism seems to be rather character-

istic and it will be remembered that luminosity measurements in the blue part of the spectrum are notoriously unreliable, even in the normal eye. The spectral sensitivity, as determined by increment threshold measurements under these conditions (i.e. with adapting field of wavelength 578 m$\mu$) is, however, shown in Fig. 4 (Stiles, 1953). The fact that this "monochromatism" produces a sensation of colour is also peculiar, since other forms of monochromatism (e.g. rod vision, central foveal vision in protanopes and deuteranopes, central foveal vision in the normal subject after adaptation to either red or blue-green, or vision by the whole eye after adaptation to strong violet light followed by adaptation to blue-green or to red) only evoke colourless or extremely unsaturated sensations. Moreover, recovery from the violet monochromatism is sudden (Brindley, 1953) as compared with that from other forms of adaptation-monochromatism; unlike most other forms of adaptation, however, it leaves a very prolonged after-image (Stiles, 1953) though, interestingly enough, adaptation to bright violet light (i.e. the complementary hue) also produces similarly prolonged after-effects.

8. After adaptation, either to strong white light or to monochromatic stimuli, the blue mechanism recovers its sensitivity at quite a different rate from the red and green mechanisms, which generally recover in comparable fashion.

After adaptation to most wavelengths, but especially to yellow and even to blue, there subsequently occurs the phenomenon of "positive blue", in which the blue mechanism appears to be abnormally prominent (Wright, 1946). By comparison with what is seen by the other (unadapted) eye, all colours are tinged with blue after such adaptation.

Stiles (1949) found that after adaptation to red light the $\pi_1$ (blue) mechanism at first gained sensitivity momentarily and then suddenly lost it to such an extent as to be virtually suppressed.

9. When a disc, whose surface is divided into black and white sectors, is rotated at a certain speed, a flickering image is obtained in which flashes of vivid bluish violet alternate with brilliant yellow. At very much higher speeds, green and pink flashes take their place. Thus, in this experiment, there are again stronger associations between blue and yellow and between green and red than between other combinations, indicating that the time relationships of excitation and recovery of the blue and yellow mechanisms are related but differ from those of the red and green mechanisms.

10. Under photopic conditions, the relationship between the critical fusion frequency (i.e. the frequency at which a flickering light fuses to give an apparently continuous sensation) and the intensity of the light is different for blue light and for red light.

11. While the blue mechanism contributes greatly to saturation and

hue, it contributes very little to luminosity; in this way it again differs from the red and green mechanisms.

12. The law of additivity of colour mixtures is found to be very nearly strictly true for wavelengths from 650 m$\mu$ to 540 m$\mu$, where the blue mechanism makes little or no contribution; but, from 540 m$\mu$ to the blue end of the spectrum, there is a progressive deviation, in the sense that the field containing the more blue component is always less bright than expected (Federov, 1958). This, again indicates that the blue mechanism contributes to luminosity in a manner which is different from that of the other two mechanisms.

13. Luminosity discrimination during violet monochromatism, as mentioned above, is very poor. It is not so noticeably poor in other forms of monochromatism, except in that associated with rod vision. In the monochromatic foveal vision of some deuteranopes it is, however, somewhat below normal.

14. The Fechner fractions† of the $\pi_1$ and $\pi_3$ mechanisms (i.e. the blue mechanisms) are much higher, at 8·7%, than that of either the $\pi_4$ and $\pi_5$ mechanisms, at 1·9% and 1·8% respectively (Stiles, 1953). It will be remembered that the $\pi_4$ and $\pi_5$ mechanisms are probably identical with the *P* (green) and *D* (red) mechanisms of the central fovea which certainly involve cones almost, if not quite, exclusively. This value of 8·7% more nearly resembles the figures associated with the rod mechanism, under scotopic conditions which are much higher than those associated with cone vision in general.

15. The quantum efficiencies ‡ of the $\pi_1$ and $\pi_3$ mechanisms ($5 \times 10^{-5}$ and $1.4 \times 10^{-5}$) are noticeably less than those of the other cone mechanisms ($\pi_4$ and $\pi_5$) which are $3.8 \times 10^{-3}$ and $3.7 \times 10^{-3}$ respectively. They are also considerably less than that of rods ($5 \times 10^{-2}$) (Barlow, 1958b). The blue mechanism therefore is, in this way, very different from the others and has a considerably lower quantum efficiency than either the rods or the cones responsible for the red and green mechanisms.

16. Measurements of summation areas§ and of visual acuity under conditions like that of violet monochromatism, in which the blue mechanism is isolated or at least highly dominant (Stiles, 1949; Brindley, 1954), suggest that the receptive fields of the cells involved in the blue mechanism are very much larger than those involved in the red and green

† The ratio of a just detectable increase in intensity to the intensity of the stimulus ($\Delta I/I$).

‡ The smallest fraction of the quanta entering the pupil needed to produce a minimal effect in the mechanism.

§ Within the summation area, intensity × area is constant, i.e. the effects of stimulation of adjacent receptors are pooled within a certain area, probably the synaptic area of bipolar cells or ganglion cells.

mechanisms, e.g. they may subtend 7 to 18′ as compared with 1 to 4′. They are somewhat smaller than those of rods in the dark-adapted state but there is evidence that rods increase their summation areas during dark-adaptation (Barlow, 1958a; Barlow, FitzHugh & Kuffler, 1957). Thus the size of the receptive field for the blue mechanism is probably comparable with the "photopic" receptive fields for rods more nearly than it is with those for the red or green mechanisms which indubitably depend upon cones.

17. When the dark-adaptation curve is plotted after the eye has been adapted to orange light, and a violet test light is used, the curve is found to have certain extra cusps on it which Auerbach & Wald (1955) have suggested may indicate the existence of violet receptors. Other explanations are, however, possible for this phenomenon, since such points could indicate the entry of some neural mechanism; e.g. the threshold of a neural pathway perhaps dependent on the activity of several types of receptor.

18. A uniform visual field, subtending about 4°, but of such low luminance as to stimulate rods only, normally appears to be colourless, but it may darken and change to blue as soon as an adjacent comparison-field stimulating mainly the cones (i.e. a red, orange or yellow field) is provided (Willmer, 1949a). A similar phenomenon has been described by Weale (1953) when a violet field ($\lambda = 460$ m$\mu$), seen in the mesopic range (i.e. in the range between scotopic (night) conditions and photopic (daylight) conditions), is surrounded by a white field. The violet increases noticeably in saturation (i.e. it becomes more intensely violet).

19. The threshold for a violet test light is raised far more by the juxtaposition of a red light than is the threshold for a similar red light by the juxtaposition of either red or violet.

20. The sensations of blue, violet and purple are often associated with twilight conditions, when rods are beginning to dominate vision (Hunt, 1952; Knowles-Middleton & Mayo, 1952). Night-blind people are sometimes tritanopic (i.e. lacking in the blue mechanism and having vision similar to that shown by the normal central fovea). Moreover, tritanopia is often associated with retinal abnormalities and particularly with those involving damage to the rods (Helmholtz, 1924; Parsons, 1924).

21. Temporary exposure to a bright source of light in a dark room generally brings on the sensations known as after-images; these are at first continuous and appear in the same colour as the stimulus; they then become discontinuous and in the complementary hue. Finally, a phase of intermittent blue and black after-images may be perceived. These intermittent blue images are not easily, if at all, evoked from the foveal region (Trezona, 1959).

22. Some curious anomalies are connected with the chromaticity of brown. Browns are, of course, not spectral colours, though a sensation of brown can be obtained from spectral orange or yellow when it is noticeably dimmer than a comparison field of the same wavelength. On the C.I.E. colour chart, browns are located, not on the spectral locus, but well within the "triangle" (see Wright, 1946) and this means that they involve the participation of the blue primary. The use of the blue primary is not, however, necessary in order to match orange or yellow, since these are both "dichromatic colours" involving only the red and the green primaries. Do these observations therefore mean that dimming the orange light automatically evokes the blue mechanism? This is certainly queer because the simple addition of blue to yellow or orange converts these colours to white or pink.

23. In some diabetics, there is a reduction of the red and green visual fields, even amounting to their total loss. This has been found to be associated with degeneration of the small ganglion cells. The blue fields, however, remain intact (Rönne, 1913). This points to a connection between the blue mechanism and the large ganglion cells which remain undamaged, and is another example of the difference between the blue mechanism and both the red and the green mechanisms.

### Rods in Photopic Vision

Such then are some of the peculiarities associated with the blue mechanism. It will have been noticed that many of them (i.e. 2–16, 18, 19, 21, 22 and 23) indicate that the blue mechanism depends on something different in kind from both the red and the green mechanisms. Moreover, several of the peculiarities strongly suggest a relationship between the blue mechanism and the rod mechanism (see 3, the retinal distribution; 13 and 14, brightness discrimination; 16, the receptive field; 18, the effects of an adjacent field on a scotopic white; 20, twilight vision; and 21, the blue after-images). It might be expected that if a "blue" cone were as different from the "red" and the "green" cones as some of its hypothetical properties suggest (see 2–8, 10–13, 18 and 19), it might be possible to differentiate this type of cone by some histological, cytochemical or developmental peculiarity or by virtue of some difference in its synaptic connections. Nevertheless, such evidence is entirely lacking. Histologically there are possibly two distinct classes of cones (Kolmer, 1936; Leach & Willmer, 1950), and there are at least two distinguishable ways in which cones connect with the bipolar cells, but beyond that there is no direct or structural evidence for the further subdivision of the cone population. It must, of course, be conceded that a cone could look exactly like the existing ones and yet have the peculiarities of function that are associated with the

blue mechanism. Meanwhile, however, there are twenty times as many rods as cones in the human eye and their function in daylight vision, if any, still eludes us.

It is very much to the point therefore to consider first how the rods are likely to behave in photopic conditions, and then to discuss whether they could possibly provide the basis for the blue mechanism.

The observations of Rushton, Campbell, Hagins & Brindley (1955), on the bleaching of the pigments in the living human eye, show quite clearly that rhodopsin is present in appreciable amounts in states of light-adaptation at which the full colour sense is in operation. Rods must therefore play some part in photopic vision unless one of the following events occurs:—

(*a*) The photochemical breakdown of rhodopsin ceases, becomes different in kind or, in some other way, becomes disconnected from the excitation of the synapses.

(*b*) The rods are so violently stimulated that they are always discharging at their maximum rate and so do not contribute differentially in different parts of the spectrum. While there is evidence that this is true at high intensities and in the spectral regions where the rods are very sensitive (Aguilar & Stiles, 1954), it would not be expected to apply at lower photopic brightness levels nor in those spectral regions where the sensitivity of the rods rapidly declines or where the sensitivity of the cones may be very similar to that of the rods. However, from experiments on animals, it is worth noting that ganglion cells connected primarily with rods do apparently attain maximum rates of discharge more easily than those to which cones connect. The saturation of the rods in photopic conditions should not, however, make their activity entirely undetectable.

(*c*) The activity of the rods is suppressed at photopic levels by a direct negative feed-back mechanism set up by the activity in neighbouring rods. It is certainly true that in animals some ganglion cells, activated only by rods, can become "silent" during illumination and that at high intensities, they function only as pure off-elements, as the result of just such a mechanism. However, once again, the lower brightness levels and the spectral extremes raise difficulties; for, under these conditions, the rods would contribute with something of their own spectral sensitivity and this is inconsistent with the colour-mixing data.

(*d*) The rhodopsin mechanism changes at higher brightnesses to become dependent upon some other photochemical substances of less sensitivity and with its maximum activity near 450 m$\mu$. Visual yellow, for example, would fill this role, but it is not known to be

photosensitive. Conceivably, the re-synthesis of rhodopsin could overtake its breakdown and this might stimulate the rod cell in an opposite sense to that achieved by the breakdown.

(*e*) As soon as the cones reach a certain threshold of activity they begin to suppress the action of the rods by direct inhibition. Several reasons can be advanced in favour of this view and the possibilities will be more fully discussed in the later parts of this paper. Some of the main reasons are as follows:—

(i) In many animals the reactions of rods and cones are known to be opposite in character; e.g. the opposite directions of their photomechanical movement in catfish, amphibia, etc. (Detwiler, 1943). The inhibition of rods by cones is, of course, not inconsistent with the existence of a simultaneous but smaller effect of the opposite kind, namely the inhibition of cones by rods. The latter effect must necessarily be smaller because cone vision certainly occurs in photopic conditions over most, if not all, of the visible spectrum.

(ii) Electrical recordings from optic nerve fibres and retinal ganglion cells show that a greater number of "on-off" units (i.e. units responding both when the light is turned on and when it is turned off) occur in retinae containing both rods and cones than in retinae containing preponderantly either rods or cones. The association of pure "on-elements" with rods and the increase in both "on-off" and "off" units when both rods and cones are present suggest inhibition by cones as well as that known to be caused by the rods themselves. Some at least of the "on-off" and "off" units have peculiarities in their spectral sensitivity curves which indicate that these units depend on two different types of receptor (Granit & Tansley, 1948).

(iii) The effects of passing constant currents directionally across the retina on the types of electrical responses obtained from the ganglion cells (Gernandt, 1947, 1948) are consistent with the view that rods and cones produce opposite electrical effects when stimulated, or that they are in some way orientated differently or polarized in opposite directions.

(iv) Some monochromats, i.e. some totally colour-blind subjects, who show central scotomata (blind spots corresponding to the foveal centre) and a probable absense of cones, also have photophobia, suggesting that normal photopic conditions are dazzlingly bright to them. This would follow if, in the normal person, the scotopic or rod mechanism becomes suppressed when the cones become active, since, in the monochromat without cones, there would be nothing to damp the activity of the rods. Man has, after all, emerged from

ancestors that were essentially nocturnal and that presumably had a highly sensitive scotopic mechanism. This would be too great an asset to lose and its temporary suppression in daylight would seem to be an obvious compromise.

(v) When a visual field, which is so dimly illuminated by light of any wavelength between 400 mμ and 580 mμ as to be stimulating rods only, has an adjacent field illuminated with a visibly yellow, orange or red light (thus stimulating cones), it is no longer seen as a colourless field but as a darker and a blue field (see p. 152(18)). The colour developed is independent of both the wavelength used for stimulation of the rods (i.e. between 400 mμ and 580 mμ) and also of that (580 mμ to 700 mμ) used for illuminating the cones (Willmer, 1949a). The darkening suggests the action of laterally-operating inhibitory processes initiated by the cones. Incidentally, but of particular relevance to a point that will be discussed later, since the field illuminated by the long-wave part of the spectrum does not overlap the area in which the rods are being stimulated, the blueness must arise, at least in part, from the stimulation of the rods for these are the only elements delimiting the size and shape of the field; but it also depends on effects spreading laterally from cones stimulated in adjacent areas in the retina.

(vi) The raising of the violet threshold by an adjacent red light (see p. 152(19)) is probably a similar phenomenon and could indicate inhibition by the cones of a mechanism particularly sensitive in the blue and violet of the region of the spectrum.

(vii) Whatever may be the exact mechanism of the inhibition of rods by cones it would operate differentially throughout the spectrum, more at the red end and less in the blue region. In daylight conditions, therefore, rods might be so effectively inhibited that they could only contribute in the blue region of the spectrum where the action of the cones would be relatively much weaker. Throughout the rest of the spectrum the rod pathway might presumably give no response at all, though the rods themselves might still be photochemically activated. The maximum sensitivity of the blue mechanism at 450 mμ and not at 500 mμ (where scotopic rods have their maximum) would thus seem to be a probable consequence of their inhibition by cones.

(viii) The inhibition of rods by cones is not inconsistent with the photochemical effect on the rods becoming at the same time saturated, or maximal, both at high brightness and in those spectral regions in which the rods are most sensitive. Nor is it inconsistent with the simultaneous occurrence of self-inhibitory processes as outlined above, nor with an opposite inhibitory action of rods on cones. All

four processes could be in operation at once and they would give to the photopic rod mechanism many properties (including spectral sensitivity) which would be different from those possessed either by the scotopic rod-pathway or by the cones.

(ix) More direct evidence for the inhibition of rods by cones comes from the study by Elenius & Heck (1957) of the changes in size of the *B* wave of the electroretinogram during the process of dark-adaptation in comparison with the increase in the visual purple content of the eye. In darkness, the latter starts to rise almost immediately, while the former shows no change at first and then a rather quick rise, which eventually subsides and thereafter coincides with the increase in visual purple. The authors interpret the sudden rise in the *B* wave after the initial delay, as being caused by the release of the rod activity, which up to that time had been suppressed by the cones. Their main reason for this interpretation is that there was no delay in the rise in the *B* wave in a subject who was totally colour-blind (i.e. a rod-monochromat), and this subject's *B* wave followed the regeneration of rhodopsin rather closely.

Since the electroretinogram, certainly in the frog and probably elsewhere also, has its origin in the immediate vicinity of the rod and cone cells (Brindley, 1958), the antagonism between the rods and cones must occur in this layer, either directly or perhaps through the intervention of horizontal cells. The pictured structure of the latter in primates (Polyak, 1941) is certainly consistent with their establishing connections between cone pedicles and rod terminals.

By comparison with the action of cells in a comparable position in fishes (see p. 144) the horizontal cells could perhaps form the basis for a blue-yellow mechanism if, when rod activity was in excess, they became electronegative, and if they became electropositive when the cones became dominant.

(x) Study of the electroretinogram under photopic conditions shows that, over much of the spectrum, the usual large "scotopic" *B* wave is very much reduced and is obscured by a quicker diphasic response (Adrian, 1945) composed of a negative *A* wave and a positive X wave, of which the latter is particularly associated with "red" cones (Armington, 1953). Nevertheless, there are traces of something which could be interpreted as a slower *B* component, particularly in the blue-green, blue and violet parts of the spectrum, even at quite high photopic levels (Armington, 1953; Johnson & Cornsweet, 1954). This component increases with increased light intensity slightly more quickly than the positive wave in the red end of the spectrum. Moreover, when, under photopic conditions, the

intensities needed to produce a given positive response are plotted against wavelength, the electroretinogram is found to be very much more sensitive in the blue region of the spectrum than would be anticipated from the standard luminosity curve. This extra sensitivity in the blue is somewhat less pronounced in two congenitally night-blind subjects, in parallel with their scotopic responses. All these factors are consistent with the view that a contribution from the rods is present in the normal photopic electroretinogram, and that it is particularly evident in the blue region. Presumably rods are still active photo-electrically, and particularly so in the region where cones are relatively inactive.

From all that has been said so far it would appear that there is a *prima facie* case for connecting the blue mechanism with rods. The other side of the picture must therefore be examined.

Three main arguments have been raised against rods participating in colour vision as one of the receptor mechanisms. The first lies in the observation that the spectral sensitivity curve for rhodopsin is not consistent with colour mixture data, for it cannot be obtained by any linear transformation of the colour mixture curves. However, since the effective response which can be made by the rods in photopic conditions may depend on (*a*) inhibition by cones, the extent of which may itself vary in a non-linear manner and only start above a certain threshold, (*b*) inhibition by the rods themselves, thus altering, in a non-linear manner, the effectiveness of any inhibition by the cones, (*c*) the possible "saturation" of the activity of the rods over quite a large part of the spectrum so that their response is uniformly maximal (apart from inhibition) over this range, (*d*) the extent to which each of the two types of cone contributes to any inhibition which they bring about, and (*e*) the possible inhibitory action of rods on cones, this argument about the shape of the sensitivity curve may not be so formidable as it appears at first sight. Rods, if they contribute to vision in photopic conditions, must clearly have their normal activity very much modified.

The second main argument which has been raised against the idea that rods participate in colour vision is embodied in the dictum that "when two colour-fields match in every respect the receptors in those fields must be stimulated to exactly the same extent in both fields". Thus, if rods form one group of those receptors, a match made under photopic conditions should still hold when the intensity is reduced to scotopic levels and the rods are then functioning alone. In practice, many trichromatic matches made under photopic conditions break down when the brightness is changed and especially under scotopic conditions. For example, the match between

two complementary whites, made respectively by mixing red with blue-green and blue with yellow, no longer holds when sufficiently dimmed.

The third main argument against the participation of rods in colour vision is the presence of the Stiles-Crawford directional effect in the blue region of the spectrum. Light falling obliquely on the cones has been found not to stimulate them so efficiently as where it falls in the direction of their long axes. This difference is not observed when light falls on rods. If therefore the rods are the receptors for the blue end of the spectrum it might be thought that the directional effect would not be observed with blue light, yet it is actually found to be present. Since, however, it is probable that for the perception of colour at least two mechanisms must be active at the time, and one of these must be cones, the difficulty does not seem to be immediately insuperable. The problem will be discussed in more detail on a later page (see p. 176) where it will be seen that the observed data are consistent with the rods being involved in the blue mechanism.

On the one hand, therefore, there is much circumstantial evidence suggesting that rods participate in colour vision and, on the other, there are at least three strong arguments against the idea. Either the rather formidable array of evidence already cited and the hypotheses to be outlined in the later part of this paper must be explained away as pure coincidence, or in some other manner, or else a satisfactory explanation must be found for such observed facts as the breakdown of photopic matches in scotopia, for the presence of a Stiles-Crawford effect in the blue, and for the failure to relate the sensitivity of the blue mechanism directly to the photosensitivity of rhodopsin.

While it cannot be pretended that these explanations are immediately forthcoming, there are several points which should be kept in mind in assessing the significance to be attached to them.

First it should be emphasized that two of the arguments are based on data obtained from colour matches. Such matches are made by the brain on the basis of information passed up to it along the optic nerves. As the result of various adaptations, summations and inhibitions, which may go on in the retina and elsewhere, the information reaching the brain is not necessarily the same as that provided by the receptors themselves. For example, inhibition exerted by the cones at the red end of the spectrum could conceivably be so intense, that over a large part of the spectrum, the rods, though physically stimulated, could make no contribution to the perceived colour. By using other pathways, e.g. those connected with reflexes, rods might still continue to contribute visual information in other ways. This would depend upon the exact seat of the inhibitory process, but from the colour point of view the inhibition of rods by cones could be

in some ways as effective in eliminating rods from contributing to matches in the long wave end of the spectrum as a pre-receptoral colour-filter could be.

Secondly, it must be emphasized that a particular level of response in a retinal nerve fibre leading from the rods could be determined by:—

1. the actual sensitivity of the rods;
2. the effects of lateral summation or inhibition by neighbouring rods;
3. the effects of inhibition by cones (probably of two sorts), including the probability that the performance of the cones is not entirely independent of that of the rods, i.e. inhibitory and summative processes may be mutual;
4. the complete "take-over" of the pathway at high intensities by the cones, as has been observed in similar systems in animals (Granit, 1949). This could be restricted to one part of the spectrum.

Thirdly, the precise meaning to be attached to the observation that two fields match needs to be very carefully considered before it can be accepted that a match necessarily means equality of stimulation of the receptors. In the following cases it clearly does not.

(i) If two separate white fields on a black ground are made to match in brightness it could be assumed, for the sake of argument, that the receptors involved by the two fields are being equally stimulated. If one of the fields is now surrounded by a white field which is brighter than itself, then the original fields will no longer match. The white surround might perhaps be expected to spread on to the original field and so to increase the brightness of the latter, but, in fact, it is this field which becomes the darker of the two; so this effect of stray light can probably be neglected. Indeed, the field which is surrounded by white may have to have its intensity raised as much as five times before a match is again achieved (Grindley, 1959). Under these conditions, therefore, two fields are now regarded as matching even though the amount of photochemical change in the receptors of one field is likely to be some five times greater than that in the other. Presumably the effects of this extra excitation are reduced by some form of lateral inhibition from receptors stimulated by the light surround. It is true that this white surround alters the conditions for the match but with ordinary conditions of colour-matching, with adjacent fields on a black ground, it seems to be at least possible that one field might well influence the effects that can be observed from the other. For example, if one side was illuminated with a colour mixture which involved a small but significant response from one of the three colour mechanisms, then the extra "inhibition" spreading from the adjacent field might well eliminate this

small response altogether and the match could then be made on the basis of the remaining two mechanisms.

(ii) In somewhat the same category as the observations just made is the observation that a yellow field of limited size can, when surrounded by blue, be made to match a similar white field not surrounded by blue. Here there is some sort of mutual interaction because, not only does the blue take the "yellowness" out of the original yellow, but the blue which changes the yellow to white may not itself be distinguishable from the black surround. One wonders how far "yellowness" in one field affects "blueness" in an adjacent field, and *vice versa*, in ordinary colour-matching procedure. The effects need not necessarily be symmetrical on a quantitative basis.

(iii) The sensation of white can be aroused in a variety of ways. Scotopic white (stimulating rods only) in one eye can be made to match photopic white (predominantly affecting cones) in the other eye. Obviously this match does not depend upon the receptors being equally stimulated on the two fields.

A small yellow field (subtending less than 30′) falling entirely on the central fovea (probably affecting two classes of receptors) can be matched with a white field in the periphery (probably affecting three types of receptors). Both appear white even though there are probably no functional blue receptors in the central fovea. It will be remembered that in ordinary colour matches yellow can be made to appear white by the addition of blue, thus presumably involving the "blue mechanism".

In red-green colour-blind subjects any wavelength affecting the cones of the central fovea, which in such subjects behaves as if it were monochromatic, can be matched both with white, as seen by the dichromatic periphery, and with a scotopic white as seen by the other eye.

The perception of white, however, should perhaps be regarded as a special case and not entirely relevant to a discussion on colour mechanisms.

(iv) On the central fovea, a red field (about 100 × threshold) can have a blue light of at least 10 × threshold added to it before becoming detectably different in hue. A similar effect is observed when red is added to blue. In terms of photosensitivity, therefore, matches may not always be very precise. Under scotopic conditions this amount of blue light might prove a very strong stimulus to the rods, and its presence or absence would be very easily detectable.

(v) Matches made with contiguous fields are easier to make, more stable and sometimes different from those made with separated fields. This difference probably indicates the prevalence of lateral effects (see (i) above) and these are almost certainly post-receptoral in origin. The critical fusion frequency of a small white field can be altered either upwards or down-

wards by the juxtaposition of another constant field which is either darker or lighter respectively than the flickering field (Graham and Granit, 1931). We do not yet know how far these lateral effects spread. They are clearly more influential on small fields than on large ones but some of them may spread for some distance and the "eye" has a habit of making fields uniform if it can.

For reasons such as the above, and in view of the arguments developed elsewhere in this paper in favour of the participation of rods in the blue mechanism, it would seem to be desirable to re-examine the concept of colour-matching on adjacent fields (i.e. standard colorimetry) rather more critically in terms of what physiological information such matches actually provide about the mechanisms involved. Perhaps some of the so-called matches, which are nearly always made on contiguous fields as the result of considered judgments with unlimited time and opportunity for adaptation and equilibration, should be re-investigated more dynamically, e.g. by reversing the fields, by flicker methods, by alternating fields, by using white surrounds and so on, in order to try to assess the parts played by direct photosensitivity, on the one hand, and by retinal integration and adaptation on the other.

The first steps in such an analysis, by more dynamic methods, have been taken by Bongard, Smirnov & Friedrich (1958) and the results are interesting. They have made matches by the method of replacement, i.e. the observer views one field after the other, on the same area of retina, and the match is regarded as complete when no change is observed on replacing one field by the other. Their experiments show that while trichromatic matches made in the periphery often break down seriously on changing the brightness level, matches made by using four primaries do not do so. One can only suppose from this that the matches made with only three primaries are not sufficiently precise and that the small unnoticed differences are magnified by changes of brightness. Moreover, since the more precise matches made with four primaries were found to be independent of brightness level, including the scotopic, it follows that the scotopic receptors must contribute to the photopic colour sense of the periphery. The rods, on numerical grounds, are likely to play a more prominent part in peripheral vision than they do in foveal vision and, in consequence, the periphery is likely to be a more sensitive testing ground for observing their effects than is the fovea. Matches made with only three primaries on central fields of less than $1 \cdot 5^\circ$ were found to be almost stable, and this may simply be that they can be made initially with the accuracy necessary to stand up to wide intensity changes.

The eye is a surprisingly precise instrument, but it is also a very versatile instrument and presumably there must be several features in its make-up

which mean compromise in performance. Although mathematical equations can be accurately applied to certain aspects of retinal performance, it should not be presumed that all the variables are sufficiently known for their application to be extended indefinitely. The fact that the law of additivity shows a progressively increasing error of as much as 14% from the red towards the blue end of the spectrum (Federov, 1958) should perhaps be a sufficient warning. There are many features of vision which indicate that the retina always tends to reduce to a minimum the messages which it transmits to the brain. It achieves this end largely by adaptive and inhibitory processes, and the resultant impulses in the optic nerve emanate mostly from those stimuli whose effects cannot be further reduced by adaptation or inhibition. Such impulses are seldom the result of the direct stimulation of the receptors only.

It may be concluded therefore that, while the arguments against the participation of rods in colour vision are certainly formidable, the evidence upon which they rest is not always above suspicion, and in view of the numerous indications that a mechanism, as different from the red and green mechanisms as is the blue mechanism, could well have its origin in the rods, it is of interest to pursue the matter further and to investigate what would happen if cones inhibited the rods on the lines which have been suggested above. The results of a tentative scheme are set out in the following pages.

### An Hypothesis concerning the Action of Rods in Photopic Vision and the Nature of the Blue Mechanism

Thus in spite of certain strong negative arguments which, though extremely formidable, need not yet be regarded as prohibitive, it seems to be worth while to explore the possibility that the blue mechanism is basically the result of the inhibition of the rods by the cones. Especially is this so since the recent observations of Hartline and his colleagues on the eye of *Limulus* (Ratliff, Miller & Hartline, 1958) have given a very clear indication of the way in which such inhibitory processes may work and of the factors which may be involved. In the eye of *Limulus* they have shown that neighbouring ommatidia (the receptor units in the compound eye) mutually inhibit each other according to a quite definite pattern which depends on the extent to which the ommatidia are each directly stimulated and on their distance apart. If $r_1$ and $r_2$ are the frequencies of impulses initiated by the two receptors when each is separately stimulated, then the final responses $R_1$ and $R_2$, when both receptors are stimulated together and after mutual inhibition has occurred, are expressed by the equations

$$R_1 = r_1 - k_1(r_2 - r_{2t}),$$

where $k_1$ is a constant depending, at least partly, on the distance between

the ommatidia and $r_{2'}$ is the threshold response necessary for exerting any inhibition at all, and

$$R_2 = r_2 - k_2(r_1 - r_{1t}),$$

where the connotations are similar.

Though in the human rods and cones the inhibition may perhaps be more properly regarded as mainly a direct antagonism, if it does not occur through the intervention of horizontal or amacrine cells, it, too, is likely to be an antagonism between the effects produced by the receptors as the result of stimulation; the magnitudes of these effects, like the impulse frequencies in *Limulus*, are likely to vary with the logarithm of the stimulus itself and with such factors as the relative numbers of the two types of receptors and their distance apart.

Thus, although the inhibition may be brought about by a different mechanism from that studied by Hartline and his colleagues in *Limulus*, it is likely to follow very much the same laws.

As a first approximation, and neglecting the possible reciprocal reduction of cone activity by the rods, no very great error is likely to be committed if the equation $B = R - K(C - C_0)$ is applied to the antagonism of rods by cones, where $B$ is the resulting response from rods, now to be considered as the basis for the "blue mechanism", $R$ is the immediate response from the rods, $C$ that of the cones and $C_0$ the threshold response at which the cones begin to inhibit the rods. In other words, the question at issue is whether it is possible, by choosing a suitable value of $K$, to derive a spectral sensitivity curve consistent with that postulated for the blue mechanism by starting with the absorption spectrum of rhodopsin, as giving a close approximation to the sensitivity of the rods, and subjecting it to an essentially subtractive process based on the spectral sensitivities of the cones, when all the curves are plotted on a logarithmic basis and an arbitrary threshold is chosen from which to measure the rod and cone responses (see Figs. 1 to 4).

If all curves of sensitivity against wavelength are plotted on a logarithmic basis and an arbitrary base line is drawn joining the limits of the visible spectrum, then the ordinates measured from this base line will, other things being equal, give a measure of the response of the units at each wavelength, and by analogy with Hartline's data these responses can be subtracted.

The first problem which arises in plotting these curves is to decide which cones shall cause the inhibition of the rods.

This difficulty can be to some extent overcome by first examining the hypothetical cases for the colour mechanisms of the protanopes and deuteranopes where either $P$ cones or $D$ cones occur in isolation or at least very heavily preponderate.

The second problem is to decide exactly how sensitive rods are with respect to cones. Flamant & Stiles (1948) have suggested on the basis of the directional sensitivity of cones and its absence in rods that, in the normal eye, cones begin to show their dominance at wavelengths longer than about 610 mμ. In Figs. 1 to 3 the sensitivity curve for rods has therefore been linked with the $P$ and $D$ curves (themselves linked at 574 mμ) at 610 mμ. The $P$ and $D$ curves are those derived from the spectral sensitivity curves of the monochromatic foveal centres of protanopes and deuteranopes and probably represent the spectral sensitivity of the two types of cone in the normal eye (Willmer, 1955).

Thirdly, it is necessary to decide where the thresholds shall be placed. Since the spectrum is normally visible from approximately 400 mμ to 700 mμ the line $T$ (Fig. 2) can be taken as a reasonable threshold from which to calculate responses, and the line $T_c$ as the threshold for the inhibitory effect of the cones.

In Fig. 2, curves showing the probable effect of $P$-cone antagonism on the response set up in the rod pathway are plotted for various values of $K$ between 1 and 2. The sensitivity curve of the $\pi_1$ mechanism (one of the "blue" mechanisms) as determined by Stiles (1953) is plotted for comparison.

It is clear from this figure that the blue mechanism of the protanope derived in this way could, according to the actual value of $K$, have a spectral sensitivity very closely approaching that which is postulated on experimental grounds, and the two mechanisms, i.e. the inhibited rods ($B_p$ mechanism) and the cones ($C_p$ mechanism) would yield the same response at about 490 mμ. This is the observed colourless or neutral point for most protanopes.

Similar treatment of the situation, but substituting the cones of the deuteranope for those of the protanope, yields the curves in Fig. 3. Stiles's $\pi_3$ mechanism (the other "blue mechanism") is plotted for comparison. Again the neutral point could well fall within the right range (see curve III) for the observed neutral point of the deuteranope, 495–510 mμ, is generally more variable than that of the protanope.

These curves, based on the data for red-green blind subjects, therefore suggest, if they can be applied to the normal eye, that the blue mechanism may normally be some resultant of both these processes, and that the $\pi_1$ and $\pi_3$ mechanisms each result from the interaction of one of the two types of cone on the rod mechanism. The average sensitivity of the blue mechanism in the normal eye probably lies somewhat between that of the protanope and that of the deuteranope.

Since the two types of cones are probably about equally stimulated at $\lambda = 578$ mμ (the colourless point in central foveal vision when only $P$

and $D$ cones are involved) the sensitivity of the blue mechanism determined after adaptation to this wavelength may give a fair indication of its spectral sensitivity in the normal eye. Fig. 4 shows Stiles's (1953) estimate for the increment threshold curve under such conditions in comparison with the mean of the curves for the "blue mechanism" of the protanopes and deuteranopes as theoretically derived above and which result when $k = 1{\cdot}75$.

It must be remembered that in all these problems the assumed logarithmic relationships between stimulus and response are only approximate and that minor modifications of the curves will almost certainly be necessary. Strongly stimulated cones probably produce relatively more inhibition, as compared with those near threshold, than the strictly logarithmic relationship would predict. Hartline's figures indicate that the amount of inhibition is closely related to the frequency of impulses and it must be remembered that this only approximates to the logarithm of the intensity of the stimulus. There is also, for example, some indication from observations on the central fovea that the effects of a given increase in stimulus to the $D$ receptors are greater than those of a corresponding increase on the $P$ receptors. Secondly the rods themselves may have reached their maximum rate of response when strongly stimulated in photopic conditions, and it is by no means certain how the antagonism of the cones would then affect the response in the residual "blue pathway". Thirdly the minimum cone response necessary to produce any inhibition or antagonism at all is uncertain. In *Limulus* a definite threshold appears to be necessary before inhibition occurs, and, obviously, if inhibition depended in any way on the frequency of impulses in the opposing cell, then there could be no inhibition till at least one impulse was propagated. In the human retina this clear-cut threshold may not be required if the antagonism is more direct, but nevertheless there probably is a threshold of some sort and, in the example given, this has been arbitrarily fixed by the line $T_C$. The arbitrary choice of $K$ to achieve a best fit takes account of some of these snags but probably not of all. For example, no account is taken of any reciprocal inhibition of cone activity resulting from stimulation of the rods.

In constructing these curves a direct and complete antagonism has been assumed. As alternatives, it may be suggested that it is possible that the cones exert just sufficient inhibition on the rods to reduce their response to equality with that of the cones themselves, or, again, that the cones effectively take over the rod pathway as soon as their response exceeds that of the inhibited rods. Some of the results obtained by Granit (1949) on the impulse frequencies from ganglion cells in the cat suggest the possibility of this latter alternative; also the spectral sensitivity of the positive wave

in the photopic electroretinogram is consistent with the idea. It is, however, more difficult to harmonize the consequences of this take-over with other data and the idea will not now be further considered, though it would effectively eliminate rods as an independent set of receptors in the red part of the spectrum, which is one of the essential conditions if they are to be considered as an integral part of the blue mechanism.

At this stage it may be permissible to digress somewhat and to consider how the sensitivity curves obtained in the manner outlined (Fig. 4) describe the colour vision of protanopes, deuteranopes and normal people. The results are of interest, though they do not help to show that the curves are necessarily the right curves.

Hecht & Shlaer (1936a and b) and Judd (1948) have investigated the nature of the colour responses of protanopes and deuteranopes and have

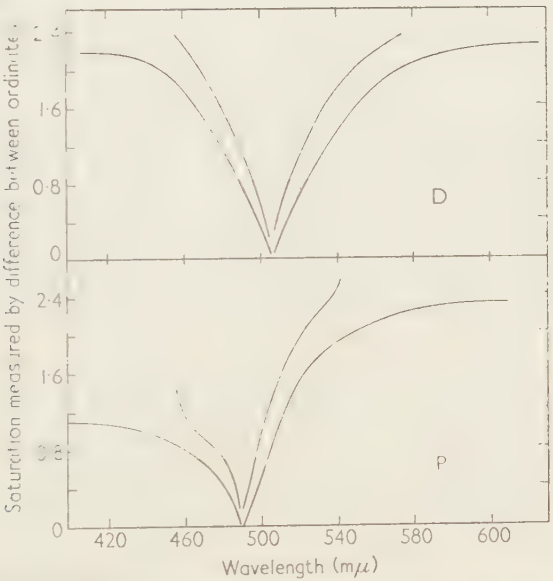



FIG. 5. Theoretical saturation of colours for deuteranopes (*D*) and for protanopes (*P*).
*D*. The difference between the ordinates of the blue mechanism (i.e. inhibited rods) and the *D* mechanism (see Fig. 3).
*P*. The differences between the ordinates of the blue mechanism and the *P* mechanism (see Fig. 2).
The curves show the limits of wavelengths within which the colours would be expected to be indistinguishable on the basis of a detectable intensity difference of 0·2 $\log_{10}$ units.

shown that subjects in both these classes see the spectrum in terms of different saturations of blue or yellow, with the change of hue occurring at a "neutral point" lying in the part of the spectrum which is blue-green to the normal eye. In the foveal centre of a normal subject a similar di-

chromatic situation occurs, giving rise to sensations of red and blue-green (König, 1894). There, the saturation of the red or blue-green colour seen at each wavelength can be effectively described by the difference between

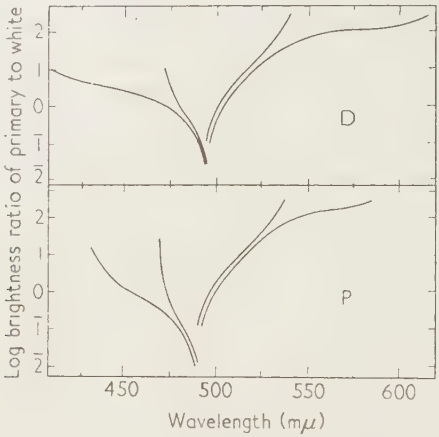



Fig. 6. Observed saturation of colours for deuteranopes (*D*) and for protanopes (*P*) (as depicted by Hecht & Shlaer, 1936b). The curves show the limits of wavelength within which the colours are indistinguishable (*cf.* Fig. 5).

the ordinates of the *P* and *D* curves at that wavelength (Willmer, 1955). With respect to the red-green blind subjects, therefore, the saturation of the spectral colours for protanopes and deuteranopes may well be described by

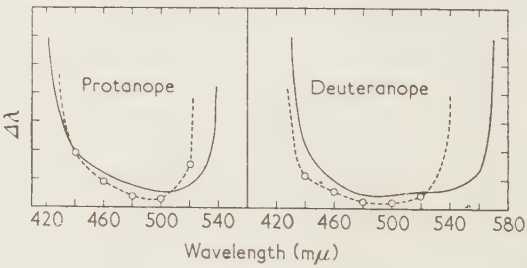



Fig. 7. Hue discrimination curves for protanopes and deuteranopes.
Continuous curves. Theoretical curves based on the rate of change of the difference between the ordinates of the *B* and *P* curves and the *B* and *D* curves respectively, with wavelength.
o – – o, hue discrimination curves as observed (data from Wright, 1946).

the differences between the ordinates of the *B* and *P* curves, and the *B* and *D* curves respectively. Such differences are plotted against wavelength in Fig. 5 and corresponding figures for the actual saturation as determined by Hecht & Shlaer (1936b) are plotted in Fig. 6. The similarity is obvious,

though the "fit" is by no means perfect. Cells affected by rods and cones in the manner of the cells in the retinae of fish, described by Svaetichin and his colleagues and mentioned earlier (see p. 144), i.e. negative-going in

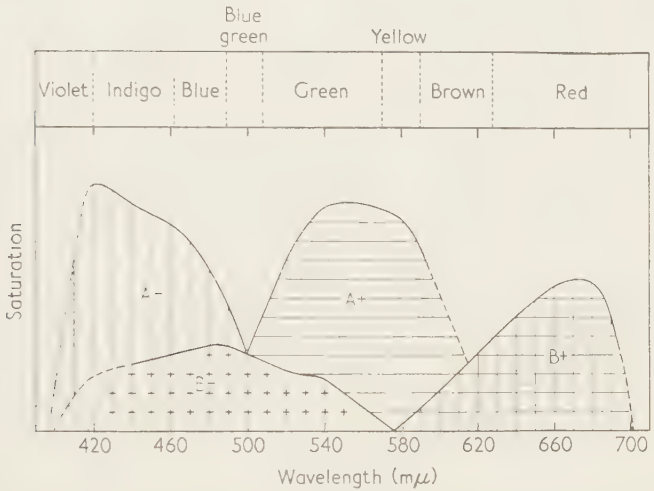



FIG. 8. Diagram to show how the activity of cells (A and B) detecting the difference between the responses of the rods and the cones, after the rods have been inhibited by the cones, could combine with similar cells detecting the difference between the responses of the *P* and *D* cones, to give the spectral colours as normally observed. The magnitude of the difference, i.e. the ordinates in this figure represent the degree of saturation of the various hues at each wavelength.

*Activity of cell* A

Blue-yellow mechanism.

A −. Vertical bars ($B - C$), giving saturation of indigo (?), the colour seen, for instance, after adaptation to intense yellow. The ordinates are measured from the lower curve as a base line.

A +. Horizontal bars ($C - B$), giving various saturations of yellow, the colour seen by deuteranopes and protanopes at the long-wave end of the spectrum. The ordinates are measured from the lower curve as base line.

*Activity of cell* B

Red-green mechanism.

B −. Crosses ($P - D$) giving various saturations of blue-green as seen in isolation at the foveal centre.

B +. Squares ($D - P$) giving various saturations of red. The blue-green response occurs in isolation at about 500 mμ and the yellow mechanism in isolation at about 574 mμ; variation in the relative sensitivities of the *P* and *D* mechanisms would account for the variable position of yellow in the spectrum as determined by different people.

It is possible that the *D* cone may be more sensitive than the *P* cone in the extreme violet thus accounting for the purple quality.

A − combined with B − yields blue
A − combined with B + yields purple
A + combined with B − yields orange
A + combined with B + yields green

response to rods and positive-going in response to cones, would seem to provide a suitable mechanism for such simple colour discrimination.

Hue discrimination for dichromats, apart from the change-over from blue to yellow, means the rate of change of saturation with wavelength or the rate of change of the difference between the ordinates of the *B* and *P* and of the *B* and *D* curves. Theoretical and observed hue discrimination curves are shown in Fig. 7. Again the similarity is striking.

If the normal eye combines both this cone *v.* rod mechanism with the *P* cone *v.* the *D* cone mechanism characteristic of the central fovea, the effects can be judged by reference to Fig. 8 where different shadings indicate the different sensations and the ordinates represent the saturation of the various sensations. It can easily be seen that the seven different spectral hues result rather satisfactorily from combinations of the information supplied by two differentiating mechanisms. The origin in this way of an apparently tetrachromatic system from what is essentially a trichromatic system is thus an intriguing method of making the Young-Helmholtz theory more acceptable to the followers of Hering and at the same time resuscitating the Ladd-Franklin and Göthlin hypotheses which have such attraction for those interested in the evolution of the colour mechanism.

The next problem which arises in connection with the blue mechanism is to explain how it contributes to luminosity and why its contribution is so small when the ordinary 1° or 2° matching fields are used.

In the paper on foveal vision (Willmer, 1955), it was shown that the central foveal luminosity curve could be derived from the *P* and *D* sensitivity curves if the contributions of each were weighted by a factor which was probably connected with the relative numbers of each type of receptor present, and if each type of receptor contributed according to the scheme illustrated on p. 6 of that paper.

By applying this principle to the three-unit system suggested above, i.e. consisting of the residual rod path (blue mechanism), the *P* path and the *D* path, it is possible to construct luminosity curves for the central fovea (30′, 50′ and 2° fields), and for fields situated 40′, 10°, 15° and 25° off-centre. These can then be compared with the luminosity curves actually recorded for these areas (Thomson & Wright, 1947; Weale, 1951, 1953). The relative numbers of rods and cones for each of these areas can be roughly estimated from the data of Østerberg replotted by Pirenne (1948). When this is done it is found that the theoretical luminosity curves obtained follow the same pattern as those actually observed; that is to say, with increasing distance from the fovea a peak of sensitivity progressively develops at about 460 mμ and a relative depression occurs at 500 mμ. However, the effect of the rods in raising this peak at increasing distances

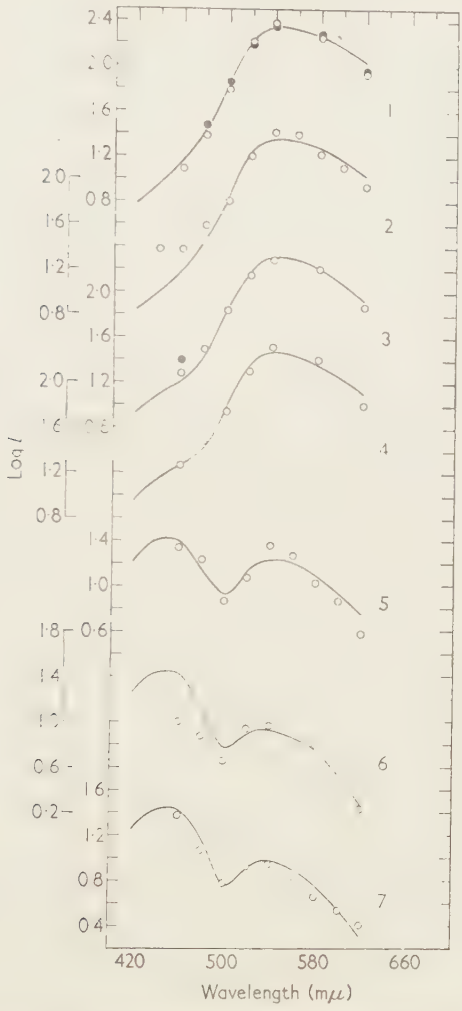



FIG. 9. Theoretical and observed luminosity curves at different retinal positions and with various sizes of field. The theoretical data are plotted according to the data given in the appendix. The results are given as the continuous curves. The open and closed circles represent the luminosities actually observed by Thomson & Wright, 1947; Weale, 1951, 1953; and Willmer, 1955.

| Curve | Centre of field | Size of field | $P/D$ | $C/B$ |
|---|---|---|---|---|
| 1 | 0° | 30′ | 2/3 | ∞ |
| 2 | 0° | 50′ | 5/6 | 100/1 |
| 3 | 0′ | 2° | 1/1 | 43/1 |
| 4 | 40′ | 15′ | 1/1 | 31/1 |
| 5 | 10° | 50′ | 2/1 | 1/1·8 |
| 6 | 15° | 50′ | 3/1 | 1/5 |
| 7 | 25° | 50′ | 4/1 | 1/4·3 |

$P/D$ = Proportion of $P$ cones to $D$ cones in the area,
$C/B$ = Proportion of cones to receptive fields for rods in the area (see Appendix).

from the fovea is greater than that required to account for the manner in which the peak actually develops on the observed curves; i.e. the inhibited rods would contribute to the luminosity curve in the manner expected but too intensely. In making this comparison the actual numbers of rods and cones, as estimated by Østerberg, were used. It is, however, arguable that the unit contributing to luminosity is not the individual rod but rather the cell which relays the effects of the stimulation of the rods, and it is possible that numerous rods impinge upon this cell. Thus the number of receptive fields for inhibited rods should perhaps be compared with the number of receptive fields for the cones and not just the number of rods with the number of cones. Since the acuity of the blue mechanism has been estimated by Brindley (1954) to be about 7·5′ while that of the red and green mechanisms is about 1′, and the Riccò areas for blue at 13′, and 4′ for red, it would seem not unreasonable to suggest that, instead of comparing the absolute numbers or rods and cones, 7·5 rods should be regarded, as a first approximation, as equivalent to one cone. In Fig. 9 are plotted the observed luminosity curves at different points in the retina together with the theoretical curves for the same areas. The latter have been constructed on the assumption that luminosity, when receptors of different sorts combine, depends more nearly on the numbers of receptive fields than on the actual numbers of receptors. Certainly the agreement between the theoretical and the observed curves is again rather striking, and suggests that the distribution of the blue mechanism in the retina follows quite closely the distribution of the rods, or rather of their receptive fields.

## Discussion

On the basis of the observations outlined above, there would thus seem to be good grounds for believing that the blue mechanism is not unconnected with the photopic activity of the rods. If rods are inhibited by cones, the two blue mechanisms, which have been demonstrated by Stiles (1953), would seem to fit with expectation since there are certainly two types of cone and each type must begin to exert its inhibitory action on the rods at an intensity level dependent on its spectral sensitivity. Thus there are at least three factors that contribute towards the sensitivity of the blue mechanism. They are:—

(i) the sensitivity of the rods themselves,
(ii) the extent of the inhibition by the *P* cones,
(iii) the extent of the inhibition by the *D* cones.

Moreover, all of these may be independently subject to change as a consequence of adaptation and mutual inhibition. Probably the nearest approach to the mean spectral sensitivity of the blue mechanism is that

obtained when the cones are both equally suppressed by adaptation to a high intensity of light of some wavelength near 574 mμ (Fig. 5). At present there is no means of knowing how effective the inhibition by the cones is, but the luminosity curves suggest that it is fairly general in all areas of the retina. The effects on the electroretinogram indicate that it occurs at a neural level near the receptors, and must therefore either be direct, or perhaps mediated by horizontal cells since Polyak's (1941) figures suggest that their structure is adapted for just such a purpose.

There are certain difficulties in the hypothesis outlined above. For example, the luminosity curves are derived by linear summation of the "responses" from "inhibited rods" and from the two types of cones, while the rods are inhibited according to Hartline's rule which involves calculating the inhibitory effects in terms of the logarithm of the sensitivity. Barlow (1958a) has shown that within a receptive field of a ganglion cell the effects of two separate stimuli summate directly and not on a logarithmic basis. This is the same effect as has been used above in calculating the luminosity curves. Thus there are precedents in eyes for both the mechanism of inhibition and for the summation postulated above, but it is not easy to see exactly how both can operate simultaneously.

Similarly it is not easy to forecast what will happen to the blue mechanism in the mesopic range, i.e. in twilight vision when the cones are barely stimulated. The extent of the inhibition must obviously decline, but whether this affects all rods alike or whether some rods, perhaps those less easily affected by cones, are completely released from inhibition before others and then contribute as ordinary visual purple receptors, as they must all do in complete scotopia, cannot yet be decided.

To the credit side of the hypothesis it may be mentioned that the different rates of adaptation in the three mechanisms seem to provide satisfactory explanations for several anomalies of the blue mechanism. For example, after adaptation to yellow the sensitivity to blue is, oddly enough, extremely strongly depressed, but it rises very rapidly in the first few seconds, much more rapidly than the sensitivity to red and green. This could be explained on the grounds that the rods of the blue mechanism are still under active inhibition by the cones for the first few seconds (see the data given by Elenius & Heck, 1957), and some release of this occurs as the cones cease to respond to the stimulus. Curiously enough, to the eye that has been adapted to yellow light or even to blue light, all colours appear to be more blue than they do to the other eye. This presumably means that while both cone responses are decreased, the blue response though made absolutely much less is changed relatively less. These phenomena are also somewhat similar to that found by Stiles (1949) in the recovery of the blue mechanism after adaptation to red light. At first the

blue mechanism appeared to be relatively little affected by the adaptation, but after a few seconds, as the red and green mechanisms recovered, the blue mechanism became even less sensitive than the green. Presumably the inhibitory action of the cones then began to assert itself.

As mentioned earlier, a sensation of brown is evoked from a spectral yellow or orange by merely dimming one half of a divided field of these colours. It certainly becomes easier to understand why the blue mechanism should become involved by dimming the yellow if dimming accentuates a lateral inhibitory process such as that suggested as necessary for the blue mechanism, which originates in the brighter (yellow) field and extends over the dimmer field.

In all the work on increment thresholds, where one field is surrounded by another, it is unfortunate that, up to the present, the nature of the observed increment is seldom clear. Detectable differences in luminosity, saturation and hue can all have different causes but any one of them could make the stimulus visible. Moreover, even if the increments are all purely luminosity increments, abrupt changes of luminosity could arise when each of the following units reaches threshold: (*a*) rods, (*b*) *P* cones, (*c*) *D* cones, (*d*) *P* cones inhibiting rods, (*e*) *D* cones inhibiting rods, (*f*) any mechanism reaching a "saturation" response. Similar considerations also must be taken into account in interpreting the "knees" on dark-adaptation curves (such as those described by Auerbach & Wald, 1955) especially after adaptation to various conditioning stimuli. The rod pathway may have its threshold altered by direct adaptation (both photochemical and neural) and also by inhibition, and there are four possible sources of that inhibition, *P* cones, *D* cones, rods themselves and some "central" source. It may therefore be possible to detect "knees" corresponding to the thresholds for *P* cones, *D* cones, rods inhibited by *P* cones, rods inhibited by *D* cones, rods inhibited by rods, in addition to the absolute threshold for the rods. Moreover these thresholds are all detected, psychophysically, by the sensory mechanism, and all or any of the "knees" found may not correspond to the receptors themselves but be determined by the thresholds of bipolar cells, ganglion cells or even of units in the visual centres themselves.

The directional sensitivity of the receptors is a phenomenon that seems to be principally connected with the cones, and rods as such are thought to be unaffected (Flamant & Stiles, 1948; Donner, 1958). Nevertheless the blue mechanism does show directional sensitivity, and this, at first sight, weighs heavily against the hypothesis that the blue receptors are rods. The situation is, however, more complicated than that because, on the above hypothesis, the cones are involved in determining the sensitivity of the rods in photopic vision when these are acting as the

basis for the blue mechanism. From the published figures for the extent of the directional sensitivity effect (Stiles, 1939) it is not easy to determine exactly how the directional effect would be expected to affect the blue mechanism. In the *dark-adapted parafovea* there is no significant directional effect from the blue end of the spectrum to about $\lambda = 600$ m$\mu$, from which wavelength the effect increases rapidly to the red end of the spectrum (see Fig. 10, Curve 1). This is probably correctly interpreted to mean that

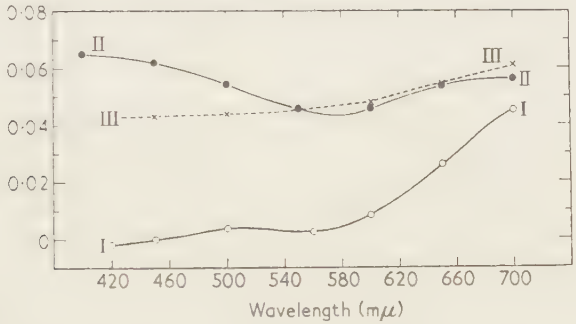



FIG. 10. The magnitude of the Stiles-Crawford effect at different wavelengths and under different conditions.

I. The dark-adapted parafovea. The rods are thought to determine the result from the blue end of the spectrum to about 600 m$\mu$ and the cones from there to the red end of the spectrum.
II. The dark-adapted fovea. The effect is predominantly on the cones.
III. The light-adapted parafovea. Cones probably predominate from about 540 m$\mu$ to 700 m$\mu$, while rods and cones together may be determining the effect from 420 m$\mu$ to 540 m$\mu$ (data from Stiles, 1939).

the rods do not show directional sensitivity, but that the cones do. In the *dark-adapted fovea*, where cones are probably involved almost exclusively, the directional effect is largest in the blue, less in the red and least in the yellow (Curve 2). On the other hand, in the *light-adapted parafovea* (unlike the dark-adapted parafovea) there is as large a directional effect in the blue (Curve 3), as in the yellow; but it is not so large as in the fovea where the "blue mechanism" is actually far more sparsely represented. This appears to be a somewhat anomalous state of affairs if the blue mechanism is only dependent on cones. If cones respond less actively to oblique rays, then their inhibitory effects on rods are presumably less when the light falls obliquely on them and thus the blue mechanism (rods inhibited by cones) should become by that amount the more sensitive. Thus, where the blue mechanism contributes in any major way to luminosity, the directional effect should be smaller than if it depended on cones only. This is exactly what is found; the directional effect on the light-

adapted parafovea, where the "blue mechanism" is more in evidence than in the fovea, is not as great in the blue as it should be if luminosity depended only on cones, as it probably does more nearly in the fovea. In the fovea the directional effect in the blue is greater than in the parafovea though the "blue mechanism" as such is less in evidence. Thus the directional sensitivity of the blue mechanism may be more complex than it appears at first sight, and it does not, of necessity, rule out the rods as being concerned in the perception of blue.

The change of hue associated with the Stiles-Crawford effect (Stiles, 1937) is also interesting and requires explanation. The changes are irregular and there are probably several causes for them. First, if cones in general are less efficiently stimulated by oblique rays, their inhibitory effect on rods is presumably less. At the blue end of the spectrum, therefore, the blue mechanism gives a greater response than with direct rays so that the colours of oblique rays must appear to match those of shorter wavelength. In the long wave end of the spectrum the blue mechanism does not enter into the system and the effect must therefore be entirely on the cones, and in order to account for the apparent shift towards shorter wavelengths, the obliquity of the light must affect the *D* cones more strongly than it does the *P* cones so that the *D* cones give a relatively smaller response, and the difference between the two is smaller. The final result of the directional effect on the appearance of colours throughout the spectrum must depend on (1) the magnitude of the effect of obliquity on the responses of the *P* and *D* cones respectively at each wavelength, and (2) the effect of this alteration of the cone response on the inhibition of the rods and consequently on the size of the residual rod response, i.e. on the response of the blue mechanism. It is, however, difficult to see why the effect at any wavelength should be of the opposite sign, as has actually been recorded for some wavelengths near 520 m$\mu$.

It thus seems to be premature to dismiss the rods from participation in photopic vision and from making their contribution to the production of the sensations of colour, and it is quite clear that the peculiarities of the blue mechanism are in urgent need of explanation.

## Appendix

The curves in Fig. 9 have been plotted in accordance with the following observations and hypotheses.

1. The ratios of the numbers of the *P* and *D* receptors are purely hypothetical and have been based on three considerations.

(*a*) The derivation of the central foveal luminosity curves from the luminosity curves of the protanope and the deuteranope indicated

the probability of a decreasing number of *D* receptors with increasing distance from the foveal centre (Willmer, 1955).

(*b*) The *D* mechanism is probably the most variable (cf. protanomaly) and is usually most concentrated in the foveal area, as indicated by the fact that the retinal field sizes for red and green are much smaller than those for blue and yellow.

(*c*) There is a possibility that the "dark cones" (Leach & Willmer, 1950) may be the *D* receptors and these become less numerous with distance from the fovea. Differences in these ratios only yield relatively small differences in the luminosity curves by comparison with the other differences discussed in this figure.

2. The ratios of *P* and *D* receptors to blue receptors are based on Østerberg's estimates (as replotted by Pirenne, 1948) of the numbers of rods and cones in one human retina. The estimated numbers of rods have been divided by a somewhat arbitrary factor of 7·5. This has been done in order to give some weight to the fact that the rods, and the blue mechanism too, have much larger receptive fields than the cones concerned with red and green. The factor of 7·5 probably errs on the small side, which means that the blue mechanism may have been given rather more influence than it actually exerts. For any more precise estimate more data are needed as to the relative numbers of rods and cones in different retinal areas and corresponding information on the sizes of receptive fields for the different mechanisms.

The following example shows the method of calculation:—

| | |
|---|---|
| Location of field | 25° |
| Size of field | 50′ |
| Ratio of rods to cones | 125/4 |
| Estimated *P*/*D* ratio | 4/1 |

1. Estimated percentage contributions

| | |
|---|---|
| *P* mechanism | 15 |
| *D* mechanism | 4 |
| *B* mechanism (rods divided by 7·5) | 81 |

2. Stimulus intensity for each mechanism at different wavelengths (Antilog. of the ordinates of the *B*, *P* and *D* curves of Fig. 4).

| | 420 | 440 | 460 | 480 | 500 | 520 | 540 | 580 | 620 mμ |
|---|---|---|---|---|---|---|---|---|---|
| *P* | 10 | 14 | 22 | 43 | 115 | 252 | 302 | 182 | 40 |
| *D* | 4 | 5 | 8 | 15 | 41 | 105 | 159 | 191 | 159 |
| *B* | 132 | 200 | 200 | 88 | 20 | 8 | 3 | 1 | 0 |

3. Weighted contributions to luminosity according to relative numbers (see 1).

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| *P* | 1·5 | 2·1 | 3·3 | 6·5 | 17 | 38 | 46 | 27 | 6 |
| *D* | 0·2 | 0·2 | 0·3 | 0·6 | 1·6 | 4 | 6 | 8 | 6 |
| *B* | 107 | 162 | 162 | 71 | 16 | 7 | 2 | 1 | 0 |

4. Total stimulus ($L$)

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 109 | 164 | 166 | 77 | 34 | 49 | 54 | 36 | 12 |

5. $\text{Log}_{10}$ total stimulus (Log $L$)

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 2·04 | 2·21 | 2·22 | 1·88 | 1·53 | 1·69 | 1·73 | 1·56 | 1·08 |

All the other curves have been plotted in an exactly similar manner.

## REFERENCES


Adrian, E. D. (1945). *J. Physiol.* **104,** 84.
Aguilar, M. & Stiles, W. S. (1954). *Optica Acta,* **1,** 59.
Armington, J. C. (1953). *J. opt. Soc. Amer.* **43,** 450.
Auerbach, E. & Wald, G. (1955). *Amer. J. Ophthal.* **39,** 24.
Barlow, H. B. (1958a). *J. Physiol.* **141,** 337.
Barlow, H. B. (1958b). National Physical Laboratory Symposium **8,** 615. H.M.S.O., London.
Barlow, H. B., Fitzhugh, R. & Kuffler, S. W. (1957). *J. Physiol.* **137,** 338.
Bongard, M. M., Smirnov, M. S. & Friedrich, L. (1958). National Physical Laboratory Symposium **8,** 325. H.M.S.O., London.
Brindley, G. S. (1953). *J. Physiol.* **121,** 332.
Brindley, G. S. (1954). *J. Physiol.* **124,** 400.
Brindley, G. S. (1958). National Physical Laboratory Symposium **8,** 509. H.M.S.O., London.
Dartnall, H. J. A. (1957). "The Visual Pigments". Methuen, London.
Detwiler, S. R. (1943). "Vertebrate Photoreceptors". Macmillan, New York.
Ditchburn, R. W. (1958). National Physical Laboratory Symposium **8,** 415. H.M.S.O., London.
Ditchburn, R. W. & Ginsborg, B. L. (1952). *Nature, Lond.* **170,** 36.
Donner, K. O. (1958). National Physical Laboratory Symposium **8,** 541. H.M.S.O., London.
Elenius, V. & Heck, J. (1957). *Nature, Lond.* **180,** 810.
Fedorov, N. T. (1958). National Physical Laboratory Symposium **8,** 305. H.M.S.O., London.
Flamant, F. & Stiles, W. S. (1948). *J. Physiol.* **107,** 187.
Gernandt, B. (1947). *J. Neurophysiol.* **10,** 303.
Gernandt, B. (1948). *Acta physiol. scand.* **15,** 286.
Graham, C. H. & Granit, R. (1931). *Amer. J. Physiol.* **98,** 664.
Granit, R. (1949). *Acta physiol. scand.* **18,** 281.
Granit, R. & Tansley, K. (1948). *J. Physiol.* **107,** 54.
Grindley, G. C. (1959). Personal communication.
Hartline, H. K. & Ratliff, F. (1957). *J. gen. Physiol.* **40,** 357.
Hartridge, H. (1947). *Phil. Trans.* **B232,** 519.
Hecht, S. & Shlaer, S. (1936a). *J. gen. Physiol.* **20,** 57.
Hecht, S. & Shlaer, S. (1936b). *J. gen. Physiol.* **20,** 83.
Helmholtz, H., von (1924). "Physiological Optics". American Optical Society.

HUNT, R. W. G. (1952). *J. opt. Soc. Amer.* **42,** 190.
JOHNSON, E. P. & CORNSWEET, T. N. (1954). *Nature, Lond.* **174,** 614.
JUDD, D. B. (1948). *J. Res. nat. Bur. Stand.* **41,** 247.
KALMUS, H. (1955). *Ann. Hum. Genet.* **20,** 39.
KNOWLES-MIDDLETON, W. E. & MAYO, E. G. (1952). *J. opt. Soc. Amer.* **42,** 116.
KOLMER, W. (1936) "Handbuch der Mikroskopischen Anatomie des Menschen, III. Die Netzhaut (Retina)", 295. (Von Mollendorf, W., ed.) Springer, Berlin.
KÖNIG, A. (1894). *S.B. preuss. Akad. Wiss.* 577.
LEACH, E. H. & WILLMER, E. N. (1950). *J. Physiol,* **111,** 48P.
MACNICHOL, E. F. JR., MACPHERSON, L. & SVAETICHIN, G. (1957). National Physical Laboratory Symposium **8,** 529. H.M.S.O., London.
PARSONS, J. H. (1924). "An Introduction to the Study of Colour Vision". Cambridge University Press, London.
PIRENNE, M. H. (1948). "Vision and the Eye". Pilot Press, London.
POLYACK, S. L. (1941). "The Retina". University of Chicago Press, Chicago.
RATLIFF, F., MILLER, W. H. & HARTLINE, H. K. (1958). *Ann. N.Y. Acad. Sci.* **74,** 210.
RIGGS, L. A., RATLIFF, F., CORNSWEET, J. C. & CORNSWEET, T. N. (1953). *J. opt. Soc. Amer.* **43,** 495.
RÖNNE, H. (1913). *v. Graefes Arch. Ophthal.* **85,** 489.
RUSHTON, W. A. H. (1957). National Physical Laboratory Symposium **8,** 71. H.M.S.O., London.
RUSHTON, W. A. H. (1958a). *Nature, Lond.* **182,** 690.
RUSHTON, W. A. H. (1958b). *Ann. N.Y. Acad. Sci.* **74,** 291.
RUSHTON, W. A. H., CAMPBELL, F. W., HAGINS, W. A. & BRINDLEY, G. S. (1955). *Optica Acta,* **1,** 183.
STILES, W. S. (1937). *Proc. roy. Soc.* **B.123,** 90.
STILES, W. S. (1939). *Proc. roy. Soc.* **B.127,** 64.
STILES, W. S. (1949). *Docum. Ophthal.* **3,** 138.
STILES, W. S. (1953). *Coloquio sobre Problemas Opticos de la Vision, Madrid.* Union Internationale de Physique pure et appliquée.
SVAETICHIN, G. & MACNICHOL, E. F. Jr. (1958). *Ann. N.Y. Acad. Sci.* **74,** 385.
THOMSON, L. C. & WRIGHT, W. D. (1947). *J. Physiol.* **105,** 316.
TREZONA, P. W. (1959). *J. Physiol.* **150,** 67.
WEALE, R. A. (1951). *J. Physiol.* **114,** 435.
WEALE, R. A. (1953). *J. Physiol.* **119,** 170.
WILLMER, E. N. (1949a). *J. Physiol.* **111,** 17P.
WILLMER, E. N. (1949b). *J. Physiol.* **110,** 422.
WILLMER, E. N. (1955). *Docum. Ophthal.* **9,** 235.
WILLMER, E. N. (1944). *Nature, Lond.* **153,** 774.
WILLMER, E. N. & WRIGHT, W. D. (1945). *Nature, Lond.* **155,** 119.
WRIGHT, W. D. (1946). "Researches on normal and defective colour vision". Kimpton, London.

*J. Theoret. Biol.* (1961) **2**, 180–189

# Electronic Aspects of the Mechanism of Action of Xanthine Oxidase†

A-M. Perault, C. Valdemoro and B. Pullman

*University of Paris, Institut de Biologie Physico-Chimique, 13, rue Pierre Curie. Paris 5e*

(*Received* 30 *October* 1960)

## 1. Introduction

The enzyme xanthine oxidase (X.O.) controls the catabolism of biological purines, in particular by metabolizing hypoxanthine and xanthine to uric acid (Bergel, Bray, Haddow & Lewin, 1957; de Renzo, 1956). Besides these two fundamental compounds, X.O. has been shown to accept as substrates a large number of heterocyclic bases: purines (Bergmann & Dikstein, 1955, 1956; Bergmann & Kwietny, 1958a, 1959; Bergmann & Ungar, 1960; Bergmann *et al.*, 1958, 1959, 1960a, b; Fridovich & Handler, 1958b; Klenow, 1952; Kwietny *et al.*, 1959; Wyngaarden, 1957; Wyngaarden & Dunn, 1957), pteridines (Bergmann & Kwietny, 1958b; Hofstee, 1949, 1956; Lowry, Bessey & Crawford, 1949), aldehydes (Blair, 1957; Booth, 1935, 1938; Knox, 1946; Mackler, Mahler & Green, 1954), etc.

In the present paper we shall try to determine the electronic aspects of the enzymic oxidation of purines. A large number of quantitative experimental data both on the pathway and the rate of oxidation of these molecules by X.O. are available, essentially through the work of F. Bergmann and his collaborators.

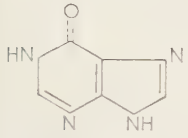
## 2. The Method

The calculations of the electronic indices of the molecular structure of purines, relevant to the problem of their enzymic oxidation, have been carried out by the L.C.A.O. approximation of the molecular orbital method of quantum chemistry (see Pullman & Pullman, 1952, for details about the method) with appropriate parameters adopted for the Coulomb and exchange integrals involving heteroatoms (Pullman & Pullman, 1958d, 1960a). A part of the data utilized in this paper originates from previously

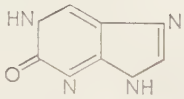
† This investigation was supported by a U.S. Public Health Service Grant No. CY-3073 (National Cancer Institute).

published calculations on the electronic structure of purines (Pullman & Pullman, 1958a, b, 1959a, 1960b; Pullman, Pullman & Berthier, 1956).

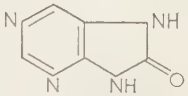
One of the major difficulties encountered in this research is related to the possibility of the existence of certain of the compounds under investigation in a number of tautomeric forms, neutral and ionic. All our calculations refer to the predominant neutral tautomeric forms of the molecules, such as manifested themselves in a series of experimental investigations



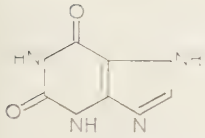

I – Hypoxanthine



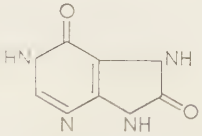

II – 2-Hydroxypurine
(2–OH–P)



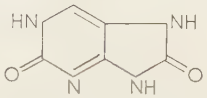

III – 8-Hydroxypurine
(8–OH–P)



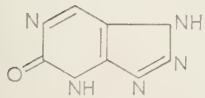

IV – Xanthine
(2, 6–OH–P)



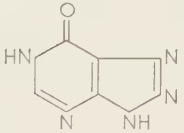

V – 6, 8-Dihydroxypurine
(6, 8–OH–P)



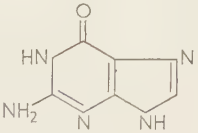

VI – 2, 8-Dihydroxypurine
(2, 8–OH–P)



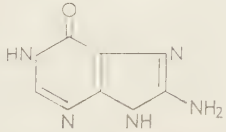

VII – 2-Hydroxy-8-azapurine
(2–OH–8–AzaP)



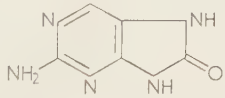

VIII – (6-Hydroxy-8-azapurine)
(6–OH–8–AzaP)




IX – Guanine
(2–$NH_2$, 6–OH–P)




X – (6-Hydroxy, 8-aminopurine)
(6–OH, 8–$NH_2$–P)




XI – (2-Amino, 8-hydroxypurine)
(2–$NH_2$, 8–OH–P)

Brown & Mason, 1957; Mason, 1954, 1957). These forms are presented n formulas I–XI. The significance of this state of affairs will be examined on a number of occasions in the discussion. For the sake of simplicity the compounds will be designated in the course of the discussion by the obvious abbreviations indicated under the name of each of them in the formulas I–XI.

One of the essential theoretical quantities involved in our discussion will be the *nucleophilic localization energy* (N.L.E.) of the carbon atoms of he purine skeleton. This denomination refers to one of the fundamental

theoretical indices, utilized in quantum chemistry for the interpretation (or prediction) of the chemical reactivity of carbons towards attacks by nucleophilic agents (e.g. negative ions). It is a measure of an important part of the activation energy for such attacks and, in a series of related molecules, it is considered as the essential variable part of the activation energy for this type of reaction. Under such conditions the relative values of this electronic index indicate the relative reactivities of the compounds or positions studied (Pullman & Pullman, 1952, 1955).

## 3. Results and Discussion

Experimentally, the oxidation of purines by X.O. takes place always on a carbon atom doubly bound to an adjacent nitrogen. It is generally supposed that the reaction takes place in two successive steps:

1. The hydration of the HC = N group.
2. The subsequent dehydrogenation of this group leading to the enol or the keto form of the oxidized purine.

The apparent complexity of this transformation makes it fairly improbable for a theoretical investigation to elucidate, at present, its detailed mechanism, at the electronic level. However, it was hoped that such an investigation might bring into evidence the particularly important electronic factors involved, at least inasmuch as the substrates are concerned (the enzyme-dependent properties being more difficult to attain). We shall consider such factors in connection, separately, with the pathway and the rate of the oxidation.

### (*i*) *The pathway of the oxidation*

In purine and a series of its monosubstituted derivatives, there exist several HC = N bonds susceptible *a priori* to be the site of the oxidation. As a rule, however, the reaction occurs essentially at a preferential position. The search for a correlation between this preferential reactivity of a given HC = N bond and its different electronic characteristics pointed to the existence of a relation between this reactivity and the value of the nucleophilic localization energy of the C atom of such bonds: *in the case of the availability of two or more HC = N bonds, the oxidized carbon is the one having the smallest nucleophilic localization energy* (the greatest ability to be attacked by nucleophilic agents such as $OH^-$ ion). This proposition is illustrated in Table 1. It can be seen that there is only one exception to our rule: the 8–OH–P which is oxidized at the C atom which has the highest N.L.E. However, this exception is not a really important one as the two available carbons have in fact very similar N.L.E.s and as the overall reactivity of this compound is relatively very small (see p. 183).

It may be noted that Table 1 does not contain the 2–OH–P. This is due to the fact that this molecule reacts essentially in the tautomeric form II, as shown by Bergmann and his colleagues in studies with its $N_1$-methylated derivative (Kwietny, Levin, Bergmann & Brown, 1959). Consequently, although it has two unsubstituted secondary carbons, it has

TABLE 1

| Compound | Available positions† | Nucleophilic localization energy (in $\beta$ units)‡ | Formal charge | Experiment references |
|---|---|---|---|---|
| 6–OH–P | 2 | 2·036 | + 0·172 | Bergmann & Dikstein, 1956 |
| | 8 | 2·349 | + 0·028 | |
| 8–OH–P | 2 | 2·376 | + 0·077 | Bergmann & Dikstein, 1956 |
| | 6 | 2·348 | + 0·031 | |
| 2–$NH_2$–P | 6 | 2·180 | + 0·098 | Bergmann, Levin & Kwietny, 1958 |
| | 8 | 2·217 | + 0·084 | |
| P | 2 | 2·323 | + 0·098 | Bergmann & Dikstein, 1956 |
| | 6 | 2·176 | + 0·093 | |
| | 8 | 2·176 | + 0·105 | |
| 8–Aza P | 2 | 2·283 | + 0·109 | Bergmann, Levin & Kwietny, 1959 |
| | 6 | 2·120 | + 0·110 | |
| 8–$NH_2$–P | 2 | 2·362 | + 0·066 | Bergmann, Kwietny, Levin & |
| | 6 | 2·238 | + 0·059 | Engelberg, 1960 |
| 6–$NH_2$–P | 2 | 2·320 | + 0·102 | Bergmann & Kwietny, 1958a; |
| | 8 | 2·246 | + 0·072 | Wyngaarden & Dunn, 1957 |
| 4–$NH_2$ pyrazolo (3,4-d)pyrimidine§ | 3 | 2·582 | − 0·047 | Feigelson, Davidson & Robins, |
| | 6 | 2·284 | + 0·117 | 1957 |

† The hydroxylation takes place at the underlined position.

‡ $\beta$ being the usual exchange integral of the molecular orbital method whose approximate value is 20 kcal/mole.

§ An isomer of adenine which is an important antitumor agent (Pullman *et al.*, 1959).

only one CH = N group susceptible to an addition reaction. In fact, the case of this molecule greatly supports the hypothesis of a hydration as an intermediate step in the action of X.O. Thus this compound undergoes the hydroxylation at its $C_8$ although the N.L.E. of this carbon ( − 2·197 $\beta$) is greater than that of its $C_6$ (− 1·840 $\beta$). But a hydration can only occur at the $N_7$–$C_8$ bond.

We have also reproduced, in the last column of Table 1, the formal $\pi$-electronic charges (Pullman & Pullman, 1959b) of the available positions. No correlation seems to exist between this index and the pathway of oxidation.

(*ii*) *The rate of the oxidation*

Although the oxidation of purines by X.O. involves a complex mechanism, composed at least of two successive steps, hydration and dehydrogenation, and although, consequently, the rate of the oxidation depends on the rates of these two reactions, the available experimental data only refer to the overall transformation. In our theoretical investigation we have tried to analyse the factors that might be responsible for the relative rate of each of the partial reactions. However, the most significant correlation that we have been able to arrive at concerns the rate of the total transformation and relates this quantity, again to the value of the N.L.E. of the position attacked in the initial step of the reaction. It appears, in fact, *that the rate of the oxidation of a purine substrate by X.O. is the greater, the smaller the N.L.E. of the carbon atom which undergoes the hydroxylation.*

This proposition is illustrated in the data collected in Tables 2, 3 and 4. The interpretation of these tables requires, however, one preliminary remark. This concerns the problem of the existence of the different purines quoted in these tables in different tautomeric forms, in particular in different states of ionization. Broadly speaking, the compounds studied here may be divided from that point of view into three groups (which correspond to the three tables):

1. Purine, aminopurines, monohydroxypurines and aminohydroxypurines; these molecules exist at pH 8·3, which is the optimum pH of the reaction, mainly in their neutral form (Albert & Brown, 1954; Brown & Mason, 1957; Mason, 1954, 1957).
2. The 8–azapurines, whose $pK_a$s are around 5 (Bendich, Giner-Sorolla & Fox, 1957; Leese & Timmis, 1958) and which consequently exist at the optimum pH in the form of anions, the proton being known to depart from the triazolo ring.
3. The dihydroxypurines, with $pK_a < 8$ (Mason, 1954) but whose anionic form is not known exactly.

Our calculations have all been performed, for technical reasons, for the neutral form of all these molecules. It is obvious, consequently, that they will be of a different degree of reliability for our present problem when applied to these different classes of purines. Naturally, they are quite applicable to the first group. In so far as the *relative* values of the N.L.E.s are concerned they also most probably reflect *the order* existing in the monoanions of the 8-azapurines, as all these compounds have $pK_a$ values of the same order of magnitude and lose their proton from the same site. On the contrary too much significance cannot be given *a priori* to the

results concerning the third group of molecules. Whatever it may be, it seems obvious that it is preferable to discuss each of these groups separately rather than all of them together.

*First group—Table* 2. The compounds may be divided, broadly speaking, into three classes: very good substrates (hypoxanthine, rate = 70, with xanthine = 100), moderate substrates (rate ≈ 20), bad substrates (rate < 2). To these three distinct ranges of rates correspond three distinct

TABLE 2

| Compound | Calculated form | Experiment: Product measured | Experiment: Initial rate % † | Experiment: N.L.E. of the oxidized C (in $\beta$ units) | Experiment references |
|---|---|---|---|---|---|
| 6–OH–P | I | Uric acid | 70 | 2·036 | Bergmann & Dikstein, 1956 |
| P | | Uric acid | 20 | 2·176 | Bergmann & Dikstein, 1956 |
| 8–$NH_2$,6–OH–P | X | 2,6–OH,8–$NH_2$–P | 18 | 2·072 | Bergmann, Kwietny, Levin & Engelberg, 1960 |
| 2–$NH_2$–P | | 2–$NH_2$,6–OH–P | 17·5 | 2·180 | Bergmann, Levin & Kwietny, 1958 |
| 2–OH–P | II | 2,8–OH–P | 16 | 2·197 | Bergmann & Dikstein, 1956 |
| 8–OH–P | III | 2,8–OH–P | 1·5 | 2·348 | Bergmann & Dikstein, 1956 |
| 8–$NH_2$–P | | 2,6–OH,8–$NH_2$–P | 1·3 | 2·238 | Bergmann, Kwietny, Levin & Engelberg, 1960 |
| 2–$NH_2$,8–OH–P | XI | 2–$NH_2$,6,8–OH–P | 1·3 | 2·344 | Bergmann, Levin & Kwietny, 1958 |
| 6–$NH_2$–P | | 2,8–OH,6–$NH_2$–P | ≃ 1 | 2·246 | Bergmann, Levin & Kwietny, 1959 |
| 2,6–$NH_2$–P | | 2,6–$NH_2$,8–OH–P | 0·5 | 2·284 | Wyngaarden, 1957 |
| 2–$NH_2$,6–OH–P | IX | 2–$NH_2$,6,8–OH–P | 0 | 2·392 | Bergmann, Levin & Kwietny, 1958; Wyngaarden, 1957 |

† Xanthine served as standard (100).

ranges of the N.L.E.s. Hypoxanthine is characterized by the smallest value of this index (2·036 $\beta$), the moderate substrates by intermediate values, comprising between 2·072 $\beta$ and 2·197 $\beta$ and the bad substrates by high values, superior to 2·246 $\beta$. Guanine which is not a substrate for X.O. has the greatest value of the N.L.E. (2·392 $\beta$). The satisfactory aspect of this correlation is particularly evident when isomers are being compared.

Thus, for example, among the three isomeric aminopurines, the 2–$NH_2$–P is a moderate substrate, comparable to purine itself, while the 8–$NH_2$ isomer and the 6–$NH_2$ isomer (adenine) are very poor substrates, a result in complete agreement with the values of the corresponding N.L.E. of the attacked carbons. The case of the three isomeric monohydroxy-purines which from the point of view of the relative rate of oxidation belong each of them to a different class of compounds (6–OH–P excellent substrate, 2–OH–P moderate substrate, 8–OH–P bad substrate) is even still more striking: the values of the N.L.E.s account entirely for this order.

*Second group—Table* 3. The order of the relative rates of oxidation of the azapurines is also satisfactorily accounted for by the values of the N.L.E.s on the appropriate C atoms. There is one exception, however, which is the 2–OH–8–azaP whose experimental reactivity is much smaller

Table 3

| Compound | Calculated form | Experiment: Product measured | Experiment: Initial rate % † | Experiment: N.L.E. of the oxidized C (in $\beta$ units) |
|---|---|---|---|---|
| 6–OH–8–AzaP | VIII | 2,6–OH–8–AzaP | 41 | 2·018 |
| 2–$NH_2$–8–AzaP | | 2–$NH_2$,6–OH–8–AzaP | 21 | 2·133 |
| 8–AzaP | | 2,6–OH–8–AzaP | 14 | 2·120 |
| 2–OH–8–AzaP | VII | 2,6–OH–8–AzaP | 6·1 | 2·036 |
| 6–$NH_2$–8–AzaP | | 2–OH,6–$NH_2$–8–AzaP | 3·1 | 2·286 |

For these experiments, see Bergmann, Levin & Kwietny, 1959.
† Xanthine served as standard (100).

than suggested by the value of the N.L.E. of its susceptible carbon. The case of this molecule is, however, exceptional from the experimental point of view too, as the sample utilized by Bergmann *et al.* (1959) was obviously hydrated before the action of X.O. started. A possible interference with the specificity of the enzymic reaction may account then for the low yield of the oxidation.

*Third group—Table* 4. This series includes two excellent substrates, xanthine and the 6,8–OH–P (both reacting at 100%) and the 2,8–OH–P which practically is not a substrate for the reaction at all.

The case of this last compound is, of course, clear, as it has no HC = N bond to be hydrated. The low value of the N.L.E. of the 6,8–OH–P is also in rather good agreement with its high rate of oxidability. On the contrary the results for xanthine are not satisfactory: the high value of the N.L.E. of its susceptible carbon is in disagreement with its excellent

behaviour as a substrate for the enzyme whose very name is xanthine oxidase. Although it is particularly disappointing to observe that among all the compounds which we have investigated, xanthine is the only real exception to our correlations, it must however be remembered that, as stated previously, this molecule is ionized at the experimental pH values

TABLE 4

| Compound | Calculated form | Experiment: Product measured | Experiment: Initial rate % | Experiment: N.L.E. of the oxidized C (in $\beta$ units) |
|---|---|---|---|---|
| 2,6–OH–P | IV | Uric acid | 100 | 2·310 |
| 6,8–OH–P | V | Uric acid | 100 | 2·066 |
| 2,8–OH–P | VI | Uric acid | 0·2 | 2·038 |

For these experiments, see Bergmann & Dikstein, 1956.

and that the structure of the anion is unknown.† Consequently, our calculations which apply to the neutral species may be far from corresponding to the actual situation.

The calculations that we have carried out, in connection with the problem studied in this paper, included the evaluation of a number of other electronic indices which seemed likely to intervene in the mechanism of the transformation involved. This included in particular the evaluation of the amount of the loss of resonance energy upon the initial hydration, of the gain in resonance energy upon the subsequent dehydrogenation, of the overall change in resonance energy upon the oxidation and of the electron-donor properties of the intermediate hydrated product. The possible significance of the resonance energy charges is, of course, obvious. The possible importance of the electron donor properties of the hydrated purines may reside in the involvement of an electron transfer in the second step of the reaction (X.O. utilizes a flavin (FAD) as a coenzyme (Fridovich & Handler, 1958a), and metal cations as activators). Following the theory developed by B. & A. Pullman (1958c) the electron donor abilities of biochemicals are determined by the energy of their highest filled molecular orbital.

All these different data are reproduced in Table 5. It can be seen that no general correlation may be observed between any of these indices and the characteristics of the reaction. Some correlations of a limited scope and of perhaps some significance, seem, however, to exist. Thus if the purines which exist at the experimental conditions in the neutral or ionized forms are considered separately and if a distinction is also made between the purines which undergo the oxidation at the pyrimidine ring and those which undergo the reaction at the imidazole ring, a partial antiparallelism is observed in each group between the rate of the oxidation and the loss of the resonance energy upon hydration of the initial base. It

† Cavalieri (Cavalieri *et al.*, 1954) attributes the first ionization of xanthine to the departure of the hydrogen from $N_3$. Ogston (1935) considers the enolization at position 6 and attributes the ionization to the deprotonation of this OH group. See also Nakajima and Pullman (1959).

may also be remarked that, although in the great majority of cases the first oxidation is accompanied by a loss of resonance energy, there is exceptionally a gain of resonance energy in the case of the oxidation of the particularly good substrates. Finally, while all the hydrated purines should be relatively good electron donors, no strict correlation seems to exist between the energy of their highest filled molecular orbital and the ease of the reaction.

TABLE 5

| Oxidized ring | Compound | Relative rate % | Difference of resonance energy between the initial and the hydrated product | Difference of resonance energy between the initial and the final product | Energy of the highest molecular orbital of the hydrated product (p.h.o.o.)† |
|---|---|---|---|---|---|
| Pyrimidine | 6–OH–8–AzaP | 41 | + 0·568 | − 0·112 | + 0·448 |
| | 2–$NH_2$–8–AzaP | 20 | + 0·713 | + 0·007 | + 0·332 |
| | 8–AzaP | 14 | + 0·770 | + 0·057 | + 0·381 |
| | 6–$NH_2$–8–AzaP | 3·1 | + 0·853 | | + 0·421 |
| | 6–OH–P | 70 | + 0·588 | − 0·079 | + 0·143 |
| | P | 20 | + 0·738 | + 0·088 | + 0·407 |
| | 6–OH,8–$NH_2$P | 18 | + 0·610 | | + 0·184 |
| | 2–$NH_2$–P | 17·5 | + 0·761 | + 0·039 | + 0·250 |
| | 8–OH–P | 1·5 | + 0·978 | + 0·179 | + 0·290 |
| | 8–$NH_2$–P | 1·3 | + 0·874 | + 0·127 | + 0·043 |
| | 2–$NH_2$,8–OH–P | 1 | + 0·924 | | + 0·162 |
| Imidazole | 2–OH–P | 16 | + 0·689 | − 0·029 | + 0·269 |
| | 6–$NH_2$–P | ~1 | + 0·700 | + 0·021 | + 0·149 |
| | 2,6–$NH_2$–P | 0·5 | + 0·781 | | + 0·183 |
| | 2–$NH_2$,6–OH–P | 0 | + 0·881 | | + 0·034 |

† In $\beta$ units.

Altogether, it seems difficult to draw any conclusion as to the importance of these limited correlations.

It appears thus that the mechanism of the enzymic oxidation of biologically important and related purines by xanthine oxidase, although undoubtedly a complex process, may be rather satisfactorily related to the value of the N.L.E. of the carbon atom on which the fixation of oxygen occurs. This state of affairs necessarily suggests that it is the initial attack by a nucleophilic agent, most probably the $OH^-$ ion, which represents the essential step of the reaction, determining both the place and the rate of the oxidation. It is hoped that the study which is under way of other categories of molecules which may function as substrates for X.O. will help to explain the role of the other factors possibly involved in the reaction.

## REFERENCES

ALBERT, A. & BROWN, D. J. (1954). *J. chem. Soc.* 2060.

BENDICH, A., GINER-SOROLLA, A. & FOX, J. J. (1957). In "The Chemistry and Biology of Purines", 3. Ciba Foundation.

BERGEL, F., BRAY, R. C., HADDOW, A. & LEWIN, I. (1957). In "The Chemistry and Biology of Purines", 256. Ciba Foundation.

BERGMANN, F. & DIKSTEIN, S. (1955). *J. Amer. chem. Soc.* **77,** 691.
BERGMANN, F. & DIKSTEIN, S. (1956). *J. biol. Chem.* **223,** 765.
BERGMANN, F. & KWIETNY, H. (1958a). *Biochim. biophys. Acta,* **28,** 100.
BERGMANN, F. & KWIETNY, H. (1958b). *Biochim. biophys. Acta,* **28,** 618.
BERGMANN, F. & KWIETNY, H. (1959). *Biochim. biophys. Acta,* **33,** 29.
BERGMANN, F. & UNGAR, H. (1960). *J. Amer. chem. Soc.* **82,** 3957.
BERGMANN, F., KWIETNY, H., LEVIN, G. & ENGELBERG, H. (1960a). *Biochim. biophys. Acta,* **37,** 433.
BERGMANN, F., KWIETNY, H., LEVIN, G. & BROWN, D. J. (1960b). *J. Amer. chem. Soc.* **82,** 598.
BERGMANN, F., LEVIN, G. & KWIETNY, H. (1958). *Biochim. biophys. Acta,* **30,** 509.
BERGMANN, F., LEVIN, G. & KWIETNY, H. (1959). *Arch. Biochem. Biophys.* **80,** 318.
BLAIR, J. A. (1957). *Biochem. J.* **65,** 209.
BOOTH, V. H. (1935). *Biochem. J.* **29,** 732.
BOOTH, V. H. (1938). *Biochem. J.* **32,** 494.
BROWN, D. J. & MASON, S. F. (1957). *J. biol. Chem.* 682.
CAVALIERI, L. F., FOX, J. J., STONE, A. & CHANG, N. (1954). *J. Amer. chem. Soc.* **76,** 1119.
DE RENZO, E. C. (1956). *Advanc. Enzymol.* **17,** 293.
FEIGELSON, P., DAVIDSON, J. D. & ROBINS, R. K. (1957). *J. biol. Chem.* **226,** 993.
FRIDOVICH, I. & HANDLER, P. (1958a). *J. biol. Chem.* **231,** 899.
FRIDOVICH, I. & HANDLER, P. (1958b). *J. biol. Chem.* **233,** 1578.
HOFSTEE, B. H. J. (1949). *J. biol. Chem.* **179,** 633.
HOFSTEE, B. H. J. (1956). *J. biol. Chem.* **216,** 235.
KLENOW, H. (1952). *Biochem. J.* **50,** 404.
KNOX, W. E. (1946). *J. biol. Chem.* **163,** 699.
KWIETNY, H., LEVIN, G., BERGMANN, F. & BROWN, D. J. (1959). *Science,* **130,** 711.
LEESE, C. L. & TIMMIS, G. M. (1958). *J. chem. Soc.* 4107.
LOWRY, D. H., BESSEY, O. A. & CRAWFORD, E. (1949). *J. biol. Chem.* **180,** 399.
MACKLER, B., MAHLER, H. R. & GREEN, D. E. (1954). *J. biol. Chem.* **210,** 149.
MASON, S. F. (1954). *J. chem. Soc.* 2071.
MASON, S. F. (1957). In "The Chemistry and Biology of Purines", 60. Ciba Foundation.
NAKAJIMA, T. & PULLMAN, B. (1959). *Bull. Soc. chim. Fr.* 663.
OGSTON, A. G. (1935). *J. chem. Soc.* 1376.
PULLMAN, A. & PULLMAN, B. (1955). "Cancérisation par les Substances chimiques et Structure moléculaire". Masson, Paris.
PULLMAN, A. & PULLMAN, B. (1958a). *Bull. Soc. chim. Fr.* 766.
PULLMAN, A. & PULLMAN, B. (1959a). *Bull. Soc. chim. Fr.* 973.
PULLMAN, A., PULLMAN, B. & BERTHIER, G. (1956). *C.R. Acad. Sci., Paris,* **243,** 380.
PULLMAN, A., PULLMAN, B. & NAKAJIMA, T. (1959). *Bull. Soc. chim. Fr.* 590.
PULLMAN, B. & PULLMAN, A. (1952). "Les Théories électroniques de la Chimie organique". Masson, Paris.
PULLMAN, B. & PULLMAN, A. (1958b). *Bull. Soc. chim. Fr.* 973.
PULLMAN, B. & PULLMAN, A. (1958c). *Proc. nat. Acad. Sci. Wash.* **44,** 1197.
PULLMAN, B. & PULLMAN, A. (1958d). "La Chimiothérapie des Cancers et les Leucémies", **201.** Centre National de Recherches Scientifiques, Paris.
PULLMAN, B. & PULLMAN, A. (1959b). *Proc. nat. Acad. Sci. Wash.* **45,** 1572.
PULLMAN, B. & PULLMAN, A. (1960a). *Rev. mod. Phys.* **32,** 428.
PULLMAN, B. & PULLMAN, A. (1960b). "Results of Quantum Mechanical Calculations of the Electronic Structure of Biochemicals". Institut de Biologie Physico-Chimique, Paris.
WYNGAARDEN, J. B. (1957). *J. biol. Chem.* **224,** 453.
WYNGAARDEN, J. B. & DUNN, J. E. (1957). *Arch. Biochem. Biophys.* **70,** 150.

*J. Theoret. Biol.* (1961) **2**, 190–203

# The Estimation of Relative Fitness of *Drosophila* Populations

P. J. Claringbold† and J. S. F. Barker‡

(*Received* 8 *November* 1960)

A mathematical model of interspecific competition is defined which may be used to describe the relative fitness of *Drosophila* populations. An outline of the numerical methods required to apply the model to experimental data is given. This model has been tested in a number of experiments, two of which are presented here as illustrative material. Log relative fitness is found to be significantly varied by some of the experimental treatments but is additive over the variables examined.

## 1. Introduction

Although the concept of relative fitness is basic in the theory of population genetics, no adequate means of determining the parameter experimentally is available. As an estimate of absolute fitness Carson (1957) uses biomass (the wet weight of hatching adults) produced per unit time in a population in equilibrium with its environment. An estimate of relative fitness termed the *competitive index* was defined by Knight and Robertson (1957) as the ratio of the numbers of a stock under test to a standard stock after one generation of competition, the initial ratio and environmental conditions being standardized. Wallace (1959) has measured relative fitness as the mean number of offspring produced by single pairs of flies in physically independent, i.e. not competing, pairs of populations maintained simultaneously.

In classical studies of simply segregating Mendelian populations, e.g. diallelic loci or chromosomal rearrangements, the concept of adaptive value relative to a defined genotype is used to explain the deviations of observed genotype frequencies from those predicted on the basis of equal absolute fitness of the three genotypes (see reviews of Dobzhansky, 1951, 1955, 1957).

This approach cannot be used in studies of the relative fitness of populations in which it is impossible to define the genotypes of individuals. A mathematical model which may be used to describe the relative fitness of two or more such populations is presented below, together with tests of

† Division of Animal Genetics, C.S.I.R.O., Sydney, Australia.
‡ Department of Animal Husbandry, University of Sydney, Sydney, Australia.

its adequacy under a variety of experimental conditions. In the work, one author (J.S.F.B.) was solely responsible for the experimentation and the other largely responsible for statistical methods.

## 2. Methods

The stocks used in the experiments were:

(i) A vermilion-eyed mutant strain (v) of *Drosophila simulans*, maintained in the laboratory by mass mating until October 1958, and since, in two population cages each averaging approximately 3,500 adults.

(ii) The wild-type Oregon-R-C strain of *Drosophila melanogaster*, maintained in this laboratory since September 1958, in a population cage averaging approximately 4,000 to 5,000 adults.

When individuals of the stocks of these two non-interbreeding species are placed together in a closed population, the outcome of the competition is invariably the elimination of *D. simulans* vermilion within a few generations. This competition has been studied in population bottles under a variety of experimental conditions. The population bottles used were an adaptation of those of Reed & Reed (1948), and have been described by Barker (1960). When a population was started, only one of the medium bottles contained medium, the other being empty. The proportions of the two species in each population were determined every two weeks, this period having been calculated as the generation interval of *D. melanogaster* under these conditions (Barker, unpublished data). Therefore, two weeks after initiation, each population was etherized and counted by eye colour, i.e. either the wild-type eye colour of *D. melanogaster* or the vermilion of *D. simulans*. After counting, the flies were returned to the central chamber of the population bottles and the empty medium bottle replaced by one containing fresh medium. Every two weeks thereafter, the older bottle was

TABLE 1

*Composition of the media used in the population bottles*

| Constituent | Medium F | Medium $S_1$ | Medium $S_2$ |
|---|---|---|---|
| Agar (g) | 8 | 10 | 10 |
| Water (ml) | 600 | 700 | 700 |
| Treacle (ml) | 80 | 80 | 80 |
| Yeast (g) | 150† | 30‡ | 30‡ |
| Cornmeal§ (g) | 80 | 80 | 80 |
| Nipagen M (g) | 1 | 2 | 4 |

† Compressed live baker's yeast (*Saccharomyces cerevisiae*)—killed by boiling during medium preparation.

‡ Inactive dried yeast powder (heat killed *Saccharomyces cerevisiae*).

§ The cornmeal used is a proprietary breakfast food.

replaced by one containing fresh medium. A medium bottle thus remained in the population bottle for four weeks.

The amount of medium in each bottle was standardized at 10 ml. The medium was seeded with live yeast using 15 g compressed yeast in 30 ml water. The amount of yeast added in this way was in some cases varied as an experimental treatment. The composition of the medium also was varied and, as it had an effect on relative fitness, should be noted (see Table 1). The population bottles were stored in a constant temperature room at 25 $\pm$ 0·5°C. This room was maintained at a constant light period of 12 hours per day (6 a.m. to 6 p.m.).

## 3. Mathematical Model

Selection between two non-interbreeding competing populations when there is no overlap of generations is discussed by Haldane (1932, p. 172). If the number of offspring in the subsequent generation per parent in the present generation is $n_S$ for population $S$ (say *D. simulans*) and $n_M = (1 - k)n_S$ for population $M$ (say *D. melanogaster*), then $k$ is the coefficient of selection in favour of population $S$ and $W_{MS} = 1 - k$ is the relative fitness of population $M$ to $S$.

In the experimental populations, the total population number per generation is not determined each generation, but the relative frequencies of the two species are determined by a census of the living adults once per generation. The census is repeated at intervals of the mean generation length. Thus if $\bar{p}_{Mi}$ is the proportion of population $M$ in generation $i$, where $i$ ranges from zero upwards, then:

$$\bar{p}_{M(i+1)}/\bar{p}_{Mi} = n_M, \tag{1}$$

$$\bar{p}_{S(i+1)}/\bar{p}_{Si} = n_S; \tag{2}$$

$$\bar{p}_{Mi} + \bar{p}_{Si} = 1. \tag{3}$$

Let
$$r_i = \bar{p}_{Mi}/\bar{p}_{Si}. \tag{4}$$

Then
$$W_{MS} = n_M/n_S, \tag{5}$$

$$= \bar{p}_{Si}\,\bar{p}_{M(i+1)}/\bar{p}_{Mi}\,\bar{p}_{S(i+1)}\ ; \tag{6}$$

$$= r_{(i+1)}/r_i, \tag{7}$$

$$= (r_i/r_0)^{1/i}. \tag{8}$$

The range of variation of relative fitness so defined is:

$$\infty > W_{MS} > 0. \tag{9}$$

An alternative scale of measurement of relative fitness is given by:

$$f_{MS} = \log W_{MS}, \tag{10}$$

with the consequences that equal fitness is defined by $f_{MS} = 0$, and the range of variation is symmetrically disposed about this point. It is clear

that $f_{MS}$ is the difference between the Malthusian parameters of the two competing species (Fisher, 1930).

If it is assumed that the interspecific competition is on the same basis as the intraspecific competition, then the distribution of the proportions of equation (4) may be taken to be the Bernoulli binomial (Aitken, 1947), and estimates of relative fitness may be obtained in a standard manner

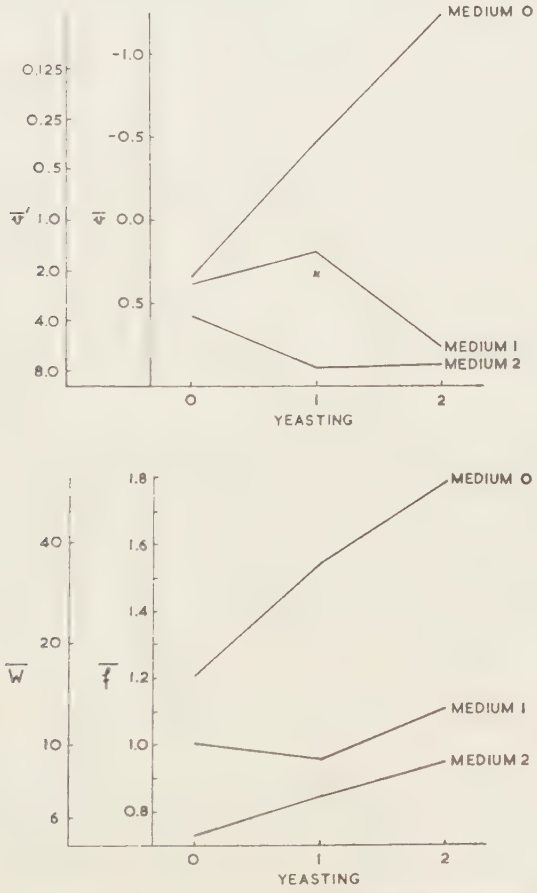



FIG. 1. Relative fitness and heterogeneity factor estimates for the medium-yeasting combinations of Experiment A. The geometric mean heterogeneity factor for Experiment B is indicated by the cross.

using the method of maximum likelihood (see Fisher, 1925; Rao, 1952). Let $p_{Mi}$ and $p_{Si}$ be the observed proportions of the species, and $n_i$ be the total number of flies, in any generation. The log likelihood $L$ of a series of observed proportions with this distribution is given in terms of the expected

proportions, themselves a function of the fitness parameter (Fisher, 1925). Differentiation of this function with respect to the parameter, and equation of the differential to zero, yields the maximum likelihood estimate of fitness, thus:

$$dL/df = \sum_{i=1}^{N} [(p_{Mi}/\bar{p}_{Mi}) - (p_{Si}/\bar{p}_{Si})]. \quad Z_i = 0 \tag{11}$$

where $$Z_i = (r_0 i W^{i-1})\,(r_0 W^i + 1)^{-2} W, \tag{12}$$

and $N$ = Number of generations scored.

The aptness of this estimate is tested by the chi-squared, Fisher (1948):

$$\chi^2_{N-1} = \sum_{i=1}^{N} (p_{Mi} - \bar{p}_{Mi})^2 n_i / \bar{p}_{Mi}\,\bar{p}_{Si}. \tag{13}$$

If there is no evidence of heterogeneity, the standard error of the estimate is given by the square root of the reciprocal of the negative of the expectation of the second derivative:

$$s^{-2}(f) = -\,\mathrm{E} \sum_{i=1}^{N} (n_i/\bar{p}_{Mi}\,\bar{p}_{Si}).Z_i^2. \tag{14}$$

On the other hand, if the chi-squared is significant, the model does not hold, and the square root of the ratio,

$$v' = \chi^2/(N-1), \tag{15}$$

or heterogeneity factor, is used to inflate the standard error of the estimate (Claringbold, 1955). If the expectations of equation (13) approach 0 or 1 such that $n_i\bar{p}_{Mi}$ or $n_i\bar{p}_{Si}$ becomes small, say less than 5, the distribution of (13) is not that of $\chi^2$, and the heterogeneity factor will have a smaller expectation than 1. This will occur if the absolute value of log relative fitness is large, unless the population size is correspondingly large (cf. Medium 0 of Fig. 1).

## 4. Numerical Analysis

Comparison of equation (11) with similar equations derived in log dose-quantal response line studies (Claringbold, 1955) reveals that the estimation problem is equivalent to fitting the logit transform to a series of proportions. With this analogy generation number is equivalent to log dose and the known initial ratio of the competing species is equivalent to a knowledge of the median effective dose. Only one parameter requires estimation, relative fitness, which is analogous to the slope of a dose-response line.

It is well known that the fitting of quantal dose-response lines is time consuming unless some simplifying assumption is made. Usually it is assumed that the probit or logit is proportional to an angle. Provided the number of observations per generation is constant or otherwise large, this

assumption considerably reduces the numerical difficulties. The assumption is only reasonable, however, if the expected proportion remains in the interval 5 to 95% (cf. Claringbold, Biggers & Emmens, 1953) and therefore cannot be applied to the present problem in which many of the observations exceed 95%. In view of the number of sets of data requiring analysis, it was decided to programme solution of the estimation equation and related equations, on an automatic computer (SILLIAC).

The log likelihood curve is bell-shaped and thus the derivative (11) when plotted against fitness forms a sigmoid curve, the central region of which crosses the abscissa with a negative slope. The equation is solved by an iterative procedure, i.e. by successive improvement of a series of values that converge towards the solution. The initial ratio is chosen so that the first generation of competition will yield a useful amount of information. The possibility of starting with very small initial ratios is limited by the total population size available with the technique. Therefore substitution of the observed proportions in equation (7) via (10) gives the first member of the series. Then equation (11) is evaluated for the complete data. Two possibilities arise, (i) the result is greater than or equal to zero, or (ii) the result is negative. In the former (latter) possibility a further member of the series is obtained by adding (subtracting) a constant to the previous value, and the differential (11) again computed. This procedure is continued until two successive members of the series yield differentials one above or on, and the other below the abscissa.

By means of the rule of false position (Whittaker and Robinson, 1944), which involves a series of either linear interpolations or extrapolations, the series of trial values rapidly converges so that the difference between the final two members may be set less than some prearranged small constant (say $10^{-3}$). The usual convergence constant (cf. Fisher & Yates, 1957) is 10% of the standard error of the estimate. Using this value some hand computers have recorded failure of convergence, presumably owing to failure of carrying enough significant figures. The constant $10^{-3}$ used in the SILLIAC programme is approximately 1% of the standard error, and in no case out of 237 sets of data did this occur; even though the initial value was very discrepant in some sets and considerable heterogeneity was present. Both of these occurrences have been suggested as the cause of failure to converge (cf. Finney, 1952).

Magnitude control was achieved in the computing programme by use of an interpretive floating point system. Per set of data the time spent in the input of information and the output of results was ($N = 3$) approximately 12 sec, while the time per iteration was approximately 1 sec. The average number of iterations was approximately 8, which could, of course, be reduced if the less stringent convergence criterion had been applied.

## 5. Experimental Evaluation of the Model

An initial series of experiments was carried out for the purposes of developing technique and testing the model under conditions of different initial ratios and population numbers. Although there appeared no systematic deviation of the observed proportions from their expectation, considerable heterogeneity was found. Such heterogeneity can only arise (i) if the model does not adequately describe the interaction between the species, or (ii) if the relative fitness varies with time due to uncontrolled environmental factors.

TABLE 2

*The design, the factors and their levels of application, of the two experiments*

| Coded factor levels | Experimental design | | | | | | |
|---|---|---|---|---|---|---|---|
| | $A = 2 \times 2 \times 3 \times 3 : 2$ | | | | $B = 2 \times 2 \times 2 : 3$ | | |
| | Factors and levels | | | | | | |
| | Starting time | Source of *D. simulans* | Medium | Yeasting (ml) | Starting time | Pre-experimental medium | Pre-experimental yeasting (ml) |
| 0 | 2 : xii : 59 | Cage 4 | $F$ | 0 | 6 : v : 60 | $F$ | 0 |
| 1 | 9 : xii : 59 | Cage 3 | $S_1$ | 0·1 | 13 : v : 60 | $S_1$ | 0·1 |
| 2 | | | $S_2$ | 0·3 | | | |

Two experiments were carried out to examine these possibilities (see Table 2). Both experiments were of the fully randomized factorial type, the replication factor being indicated as the last term in the "equation" defining the experiment. In each experiment one half of the populations were set up on the one day, the other half, one week later. Two separate source populations of *D. simulans* v, which had been maintained in population cages, were used in Experiment A. In Experiment B the possible influence of the medium and yeasting on which the *D. simulans* v flies for the initial populations had developed is studied. The results of the two experiments, in terms of the mean values of relative fitness, log relative fitness, and heterogeneity factor for each treatment combination are given

in Tables 3 and 4. The detailed results of these and other experiments will be presented in a later paper. However, it may be seen that the fitness of *melanogaster* (Oregon-R-C) relative to that of *simulans* v is quite high, so that in most of the populations of these experiments the latter species is eliminated by the third generation.

TABLE 3

*Mean values of relative fitness* ($\overline{W}$), *log relative fitness* ($\bar{f}$), *and geometric mean heterogeneity factor* ($\bar{v}'$) *for each experimental factor combination in Experiment A*†

| Experimental factors and their coded levels | | | | $\overline{W}$ | $\bar{f}$ | $\bar{v}'$ |
|---|---|---|---|---|---|---|
| Starting time | Source of *D. simulans* | Medium | Yeasting | | | |
| 0 | 0 | 0 | 0 | 48·60 | 1·55 | 1·59 |
| | | | 1 | 81·65 | 1·91 | 0·30 |
| | | | 2 | 73·10 | 1·83 | 0·06 |
| | | 1 | 0 | 27·22 | 1·43 | 2·30 |
| | | | 1 | 19·61 | 1·27 | 2·25 |
| | | | 2 | 29·14 | 1·33 | 2·49 |
| | | 2 | 0 | 5·78 | 0·76 | 5·30 |
| | | | 1 | 13·31 | 1·12 | 8·60 |
| | | | 2 | 16·55 | 1·17 | 2·54 |
| | 1 | 0 | 0 | 30·80 | 1·49 | 0·44 |
| | | | 1 | 35·99 | 1·54 | 0·99 |
| | | | 2 | 75·07 | 1·87 | 0·01 |
| | | 1 | 0 | 12·72 | 1·04 | 4·19 |
| | | | 1 | 15·20 | 0·94 | 0·67 |
| | | | 2 | 24·51 | 1·39 | 3·31 |
| | | 2 | 0 | 8·69 | 0·93 | 0·97 |
| | | | 1 | 8·92 | 0·94 | 9·41 |
| | | | 2 | 12·16 | 1·08 | 6·03 |
| 1 | 0 | 0 | 0 | 10·82 | 0·97 | 1·77 |
| | | | 1 | 29·69 | 1·47 | 0·11 |
| | | | 2 | 345·45 | 1·92 | 0·08 |
| | | 1 | 0 | 5·64 | 0·75 | 0·93 |
| | | | 1 | 6·97 | 0·84 | 3·14 |
| | | | 2 | 7·10 | 0·85 | 18·01 |
| | | 2 | 0 | 4·68 | 0·67 | 3·97 |
| | | | 1 | 4·11 | 0·61 | 7·75 |
| | | | 2 | 5·68 | 0·75 | 21·98 |
| | 1 | 0 | 0 | 6·60 | 0·82 | 19·19 |
| | | | 1 | 19·57 | 1·28 | 0·35 |
| | | | 2 | 34·72 | 1·54 | 0·23 |
| | | 1 | 0 | 7·04 | 0·84 | 3·66 |
| | | | 1 | 5·95 | 0·77 | 1·31 |
| | | | 2 | 8·73 | 0·87 | 8·23 |
| | | 2 | 0 | 3·70 | 0·56 | 9·80 |
| | | | 1 | 5·53 | 0·72 | 6·35 |
| | | | 2 | 6·26 | 0·80 | 8·03 |

† $\bar{f}$ does not equal $\log_{10}\overline{W}$, as $\bar{f}$ and $\overline{W}$ are each means of the estimates for the replicates in each experimental factor combination.

TABLE 4

*Mean values of relative fitness ($\overline{W}$), log relative fitness ($\bar{f}$), and geometric mean heterogeneity factor ($\bar{v}'$) for each experimental factor combination in Experiment B*†

| Experimental factors and their coded levels | | | $\overline{W}$ | $\bar{f}$ | $\bar{v}'$ |
|---|---|---|---|---|---|
| Starting time | Pre-experimental medium | Pre-experimental yeasting | | | |
| 0 | 0 | 0 | 21·53 | 1·32 | 1·30 |
| | | 1 | 15·31 | 1·16 | 3·42 |
| | 1 | 0 | 66·56 | 1·63 | 0·58 |
| | | 1 | 43·35 | 1·59 | 2·88 |
| 1 | 0 | 0 | 9·55 | 0·99 | 3·77 |
| | | 1 | 18·70 | 1·17 | 1·51 |
| | 1 | 0 | 8·41 | 0·90 | 4·75 |
| | | 1 | 10·34 | 1·00 | 2·30 |

† $\bar{f}$ does not equal $\log_{10}\overline{W}$, as $\bar{f}$ and $\overline{W}$ are each means of the estimates for the replicates in each experimental factor combination.

TABLE 5

*Analysis of variance of the data of Experiment A*

| Source of variation | D.F. | Standardized mean squares or product | | | | |
|---|---|---|---|---|---|---|
| | | $f^2$ | $v^2$ | fv | $(f-v)^2$ | $(f+v)^2$ |
| Time | 1 | 24·742† | 2·267 | − 7·488 | 41·985† | 12·033‡ |
| Stock | 1 | 1·814 | 0·013 | − 0·149 | 2·125 | 1·529 |
| Medium | (2) | | | | | |
| 0 v. 2 (a) | 1 | 56·165† | 27·218† | − 39·106 | 161·595† | 5·171§ |
| 0 + 2 v. 1 (b) | 1 | 3·804 | 2·013 | − 2·769 | 11·355 | 0·279 |
| Yeasting | (2) | | | | | |
| 0 v. 2 (a) | 1 | 11·186‡ | 1·658 | − 4·306 | 21·456§ | 4·232§ |
| 0 + 2 v. 1 (b) | 1 | 0·031 | 0·129 | 0·068 | 0·024 | 0·296 |
| *Interactions* | (29) | | | | | |
| Medium × yeasting | (4) | | | | | |
| a × a | 1 | 2·732 | 10·378‡ | − 5·322 | 23·754‡ | 2·466 |
| a × b | 1 | 0·052 | 0·173 | − 0·095 | 0·415 | 0·035 |
| b × a | 1 | 2·639 | 4·316§ | − 3·378 | 13·711§ | 0·199 |
| b × b | 1 | 0·649 | 1·084 | 0·840 | 0·053 | 3·413 |
| Remainder | 25 | 0·371 | 0·480 | − 0·115 | 1·081 | 0·621 |
| Error | 36 | 1·000 | 1·000 | − 0·514 | 3·028 | 0·972 |
| | | $s_f =$ 0·311 | $s_v =$ 0·474 | $r_{fv} =$ − 0·514‡ | | |

† $P < 0{\cdot}001$. ‡ $P < 0{\cdot}01$. § $P < 0{\cdot}05$.

TABLE 6

*Analysis of variance of the data of Experiment B*

| Source of variation | D.F. | Standardized mean squares or product | | | | |
|---|---|---|---|---|---|---|
| | | $f^2$ | $v^2$ | fv | $(f-v)^2$ | $(f+v)^2$ |
| Time | 1 | 15·242† | 0·625 | − 3·093 | 22·053‡ | 9·681‡ |
| Medium | 1 | 1·258 | 0·609 | 0·877 | 0·113 | 3·621 |
| Yeasting | 1 | 0·045 | 0·016 | − 0·031 | 0·123 | − 0·001 |
| *Interactions* | | | | | | |
| Time × medium | 1 | 5·682‡ | 1·406 | − 2·831 | 12·750‡ | 1·426 |
| Time × yeasting | 1 | 1·379 | 1·063 | − 1·216 | 4·874 | 0·010 |
| Medium × yeasting | 1 | 0·000 | 0·094 | 0·015 | 0·064 | 0·124 |
| Time × medium × yeasting | 1 | 0·212 | 0·016 | 0·062 | 0·104 | 0·352 |
| Error | 16 | 1·000 | 1·000 | − 0·385 | 2·770 | 1·230 |
| | | $s_f =$ 0·812 | $s_v =$ 0·800 | $r_{fv} =$ − 0·385 | | |

† $P < 0{\cdot}01$. ‡ $P < 0{\cdot}05$.

For every experimental population an estimate of log relative fitness ($f$), and of log heterogeneity factor ($v = \log v'$, see equation 15), was computed. The log transform was applied to the heterogeneity factor since it renders the variance of the measure independent of its magnitude (Rao, 1952, p. 225). An analysis of covariance of these variates was made, see Tables 5 and 6. In both experiments replication provided an independent estimate of the error variance and covariance. The significance of all interaction terms was examined against this error, and the non-significant items were pooled into a residual term (in Experiment A) merely to shorten the table.

In both analyses a negative correlation coefficient is found between log relative fitness and log heterogeneity factor. Since the difference between the two estimated coefficients is not significant, the weighted mean coefficient obtained by transformation to $z$ (see Fisher, 1948, p. 203) is calculated,

$$\begin{aligned} z' - z'' &= 0{\cdot}16 \pm 0{\cdot}33, \\ \bar{z} &= -0{\cdot}52 \pm 0{\cdot}15, \\ \bar{r} &= -0{\cdot}48. \end{aligned} \tag{16}$$

A correlation coefficient of approximately − 0·50 implies that the ratio of the major to the minor axis of the correlation ellipse is 3 : 1, viz.

$$\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}. \tag{17}$$

The direction of the major axis is given as the difference between the standardized variates. Further analyses of variance were therefore constructed on this basis, see Tables 5 and 6, and it is seen that there is less pronounced treatment variation in the direction of the minor axis of the ellipse. The relationship between log relative fitness and log heterogeneity factor is maintained over all combinations of the experimental variates (see Fig. 1).

The analysis of variance of log relative fitness values indicates that the parameter is strongly varied by secular changes, and properties of the medium in Experiment A. The absence of any significant interactions with the one exception at the 5% level in Experiment B, out of a total of 33 single degree of freedom tests, indicates that the log relative fitness measure is additive over the variables examined.

The results of Experiment B, in which media and yeasting conditions were those coded 1 in Experiment A, were uniformly heterogeneous (a geometric mean $\bar{v}' = 2{\cdot}16$), that is, approximately the value obtained in Experiment A, central point, which indicates a consistency of results.

## 6. Discussion

It has been pointed out above that heterogeneity can only arise either if the model does not adequately describe the interaction between the species, or if the relative fitnesses vary with time due to uncontrolled environmental factors. If the former is correct, then obviously this model cannot validly be used for the estimation of relative fitness. However, if the model does hold but uncontrolled environmental factors give rise to the heterogeneity observed, then it should be possible to improve the experimental technique so that this heterogeneity is removed.

The fitness of a particular *D. melanogaster* population is described by the model presented here as a single parameter, fitness relative to that of a standard *D. simulans* v population taken as one. To understand the superior relative fitness of *melanogaster* and to analyse the interaction between the two species, the various components of fitness must be considered, for example, (i) fecundity, (ii) hatchability, (iii) rate of development of immature stages, (iv) the numbers of adults emerging, (v) longevity of the adults. It should be pointed out that various attempts have been made to estimate relative fitnesses of genotypes by combining estimates of various of these components (e.g. Wallace, 1948; Buzzati-Traverso, 1955; Dobzhansky & Levene, 1955; Spiess & Schuellein, 1956; and Spiess, 1958). This technique for the estimation of relative fitness is particularly tedious and, although it may be satisfactory in some cases (e.g. Spiess, 1958), its potential dangers should not be overlooked (Barker, 1958; see also Dobzhansky, 1956). The rate of development of immature stages at

25°C is slightly faster for *melanogaster* than for *simulans* (Moore, 1952; Herskowitz, 1953), while the mean generation interval of *melanogaster* (Oregon-R-C) is two to four days shorter than that of *simulans* v (Barker, unpublished). As the populations were counted at intervals equal to the mean generation interval of *melanogaster* (Oregon-R-C) it may be thought that this will give a consistent bias in favour of this species, that is, that the relative fitness of *melanogaster* will increase proportionately with time. However, it should be noted that it is relative fitness that is being considered, and the generation interval is a component of this parameter.

The assumption of the Bernoulli binomial distribution in defining the random aspect of the interspecific competition in effect means that it is assumed that intraspecific and interspecific competition are randomized. Moore (1952) observed, however, that oviposition sites differed in the two species. Although both species found the edges of the medium surface the most desirable, it was found that *simulans* would deposit eggs in the centre more readily than *melanogaster*. At pupation also, this tendency to localize differentially has been observed in these experiments and by Herskowitz (1953). Larvae of *melanogaster* generally move on to the glass walls of the medium bottle or into the polythene central chamber of the population bottle in order to pupate, while *simulans* larvae tend to pupate on the surface of the medium. This fact, in addition to the difference in mean generation interval, probably accounts for much of the difference in fitness between the two species. The pupal survival of *simulans* is probably less than that of *melanogaster* because, as the medium becomes soft from the larval activity, many of the pupae on the surface are either submerged in the medium or displaced so that successful eclosion is made more difficult, if not impossible. It might be thought that such differentiation in oviposition and pupation sites could well lead to heterogeneity because of non-randomness in the competition. However, once the eggs hatch the larvae are very active and can range throughout the medium. Further, neither the eggs nor the pupae are stages in the life cycle involved in competition, either intra- or interspecific.

As in the populations where heterogeneity was significant, there was no suggestion of systematic deviations of the observed proportions from the expected, and as there is no evidence for any factor that would give rise to non-randomness in the interspecific competition, it would seem that the model does adequately describe the interaction between the species. It must be concluded that the heterogeneity observed is due to the fitness of the two species varying with time due to uncontrolled environmental conditions.

An explanation of the negative correlation between log relative fitness and log heterogeneity factor is simple. When relative fitness is high (say

$> 30$), the proportion of *melanogaster* in the population reaches 80% or more in one generation. Most of the information on relative fitness is thus contained in this first generation and it is consequently difficult to detect any heterogeneity. However, if relative fitness is low, the increase in the proportion of *melanogaster* is much slower, so that considerable information on relative fitness is obtained from all three generations. Any fluctuation in fitness with time will then be more easily detected. Therefore, for best use of this model, experimental conditions, the standard population of *simulans*, and the initial ratio should be chosen so that a number of generations (if possible) give proportions in the range 10 to 90%.

It has been shown therefore that this model may be used to estimate the relative fitness of different *D. melanogaster* populations. However, as relative fitness is significantly affected by the experimental conditions, any estimates so obtained are valid only for the particular set of conditions used. Nevertheless, unless there are interactions between *D. melanogaster* population and experimental conditions, the ratio of log relative fitness of one population to that of another, as determined for a particular set of experimental conditions, will hold over any set of conditions.

We are indebted to Dorothy Allingham, Gillian Davey and Robin Johnston for their assistance in the maintenance and scoring of the population bottles. This study has been supported by a University Research Grant to one of us (J.S.F.B.).

## REFERENCES

Aitken, A. C. (1947). "Statistical Mathematics". Oliver and Boyd, Edinburgh.
Barker, J. S. F. (1958). *Aust. J. biol Sci.* **11**, 613.
Barker, J. S. F. (1960). *Drosophila Inform. Serv.* **34**. In press.
Buzzati-Traverso, A. A. (1955). *Heredity*, **9**, 153.
Carson, H. L. (1957). *Genetics*, **42**, 363.
Claringbold, P. J. (1955). *Biometrics*, **11**, 481.
Claringbold, P. J., Biggers, J. D. & Emmens, C. W. (1953). *Biometrics*, **9**, 467.
Dobzhansky, Th. (1951). "Genetics and the Origin of Species". Columbia University Press, New York.
Dobzhansky, Th. (1955). *Cold Spr. Harb. Symp. quant. Biol.* **20**, 1.
Dobzhansky, Th. (1956). *Amer. Nat.* **90**, 337.
Dobzhansky, Th. (1957). *Cold Spr. Harb. Symp. quant. Biol.* **22**, 385.
Dobzhansky, Th. & Levene, H. (1955). *Genetics*, **40**, 797.
Finney, D. J. (1952). "Probit Analysis". Cambridge University Press, London.
Fisher, R. A. (1925). *Proc. Camb. phil. Soc.* **22**, 700.
Fisher, R. A. (1930). "The Genetical Theory of Natural Selection". Oxford University Press, London.
Fisher, R. A. (1948). "Statistical Methods for Research Workers". Oliver and Boyd, Edinburgh.
Fisher, R. A. & Yates, F. (1957). "Statistical Tables for Biological, Agricultural and Medical Research". Oliver and Boyd, Edinburgh.
Haldane, J. B. S. (1932). "The Causes of Evolution". Longmans, Green & Co., London.

HERSKOWITZ, I. H. (1953). *Amer. Nat.* **87,** 113.
KNIGHT, G. R. & ROBERTSON, A. (1957). *Genetics,* **42,** 524.
MOORE, J. A. (1952). *Evolution,* **6,** 407.
RAO, C. R. (1952). "Advanced Statistical Methods in Biometric Research". John Wiley and Sons, Inc., New York.
REED, S. C. & REED, E. W. (1948). *Evolution,* **2,** 176.
SPIESS, E. B. (1958). *Evolution,* **12,** 234.
SPIESS, E. B. & SCHUELLEIN, R. J. (1956). *Genetics,* **41,** 501.
WALLACE, B. (1948). *Evolution,* **2,** 189.
WALLACE, B. (1959). *Amer. Nat.* **93,** 295.
WHITTAKER, E. & ROBINSON, G. (1944). "The Calculus of Observations". Blackie and Son Ltd., London.

*J. Theoret. Biol.* (1961) **2**, 204–235.

# Outline of a Theory of Thought-Processes and Thinking Machines

E. R. Caianiello

*Istituto di Fisica Teorica dell'Università di Napoli*
*(Sottosezione di Napoli dell'Istituto Nazionale di Fisica Nucleare)*

(*Received* 9 *December* 1960)

Thought-processes and certain typical mental phenomena are schematized into exact mathematical definitions, in terms of a theory which, with the assumption that learning is a relatively slow process, reduces to two sets of equations: "neuronic equations", with fixed coefficients, which determine the instantaneous behavior, "mnemonic equations", which determine the long-term behavior of a "model of the brain" or "thinking machine". A qualitative but rigorous discussion shows that this machine exhibits, as a necessary consequence of the theory, many properties that are typical of the living brain: including need to "sleep", ability spontaneously to form new ideas (patterns) which associate old ones, self-organization towards more reliable operation, and many others. Future works will deal with the quantitative solution of these equations and with concrete problems of construction—things that appear reasonably feasible. With a transposition of names, this theory could be applied to many sorts of social or, more generally, "collective" problems.

## 1. Introduction

### A. LEVELS OF APPROACH

Attempts at a quantitative understanding and analysis of thought-processes, with or without the explicit aim of devising machines that should reproduce functions typical of the living nervous system, date as far back as Ramon Lull's syllogistic wheels. They have become a recognized and major part of scientific investigation since N. Wiener's celebrated enunciation of the principles of Cybernetics; herein lies indeed clearly, much more than in specialized studies of circuitry or of information theory, the heart and scope of this new science, which aims at synthesis as well as analysis.

The investigation of the mechanism of thought has been undertaken with a variety of methods, ranging e.g. from the study of systems that should mechanize the operations of Aristotelian logic without any requirement of similarity to living structures, to the faithful electronic reproduction of populations of hundreds or thousands of neurons. We shall benefit

from all these discussions in that they permit us to reduce the verbal presentation of our own concepts to a bare minimum, since they have made abundantly clear with what cautions and restrictions one should accept for example the very expression "mechanical thought"; otherwise we shall restrict our treatment exclusively to the presentation of our approach to this problem, as we feel that in such a field judgment is passed better *a posteriori* than *a priori*, on the ground of concrete results—which are yet to be borne by any theory, including ours—than of mere opinion.

The present outline of a theory of thought-processes is the result of about three years of discussions with people who have been working with the same premises in various fields of neuroanatomy, mathematics and theoretical physics; it also reflects, of course, the evolution of our own ideas through many discussions with guests and hosts. Our main guiding principle has been the conviction, strengthened by these discussions, that the human brain, tremendous in its complexity, yet obeys, if one looks at the operation of individual neurons, dynamical laws that are not necessarily complicated; and that these laws are such as to engender in large neuronal assemblies collective modes of behavior, to which thought-processes are correlated. A convenient formulation of these laws appears therefore as the primary objective of a research of this nature; it can only be achieved by trial and error by the process, familiar in the physical sciences, of abstracting what seems relevant into a simplified model of the real thing. The present work is one such trial; its novelty is not, of course, in the concepts just mentioned, which are as old as physics itself, although they have not yet gained general acceptance among neurophysiologists, but in the attempt made here to give them a precise and quantitative formulation.

Constant resort to neuroanatomy and neurophysiology, which is the keystone of our approach, appears necessary at two different levels: the "elementary level", which studies the individual neurons and the connections, or synapses, between neurons; the "integrative level", which studies the structure and function of specially connected assemblies of neurons, which may act as a whole and play in the nervous system a rôle similar to that of specialized organs in the body. The integrative level compares with the first as the physics of matter does with that of the atom, and is of course as essential to the understanding of the functions of a brain or of a thinking machine; we firmly believe, against the opposite views which we have heard expressed, that a study at the elementary level is as essential to the second as one type of physics is to the other.

We shall have very little to say here about the integrative approach, in which many more investigations are needed before a satisfactory state of knowledge is achieved, except that our equations, once the appropriate connections among neurons of a given assembly are introduced into them,

will permit the quantitative study of its collective behavior as a whole. It is our belief that the subsystems of a brain are quite different, in structure and complication, from the standard circuits of electronics, and that there will be a great deal to learn in this respect from neuroanatomy; also, that a thinking machine built for some special purpose may well need organs, or subsystems, organized quite differently from those of the animal brain, although the same elementary laws will be valid. Our equations are also intended to provide a useful tool both for theoretical study and for experimentation in this respect.

Because of the lack of definite knowledge and of general agreement among specialists on many facts of neuroanatomy and neurophysiology on one hand, and of the great wealth of available observational material on the other, we think it best to present our views as the direct description, reduced to bare essentials, of a *model* of the brain forgoing the detailed analysis of anatomical data from which, in fact, our considerations stem.

### B. THE MODEL

By "model" or "machine" we mean exclusively a device *that can actually be built*, and which operates according to mathematical equations that are *exactly known and numerically solvable* to any wanted accuracy. Although this necessarily implies drastic schematizations and simplifications, it is hoped that the features essential to thought-production are retained by the model; successive approximations to reality will require improvements in the structure of the machine and in its operational laws, but at each step one must know exactly what is being done. Without a complete mathematical control of the situation, a machine may perhaps think, but one would hardly know why or how.

Mathematically, our model consists of two sets of equations: the "neuronic equations", which describe the instantaneous operation of the machine; the "mnemonic equations", which describe the growth of memory into it. From these equations it is possible to predict and study the "mental" phenomena which are typical of such a machine: learning, forgetting, re-integration, conditioning, analysis of patterns and spontaneous formation of new patterns, self-organization into reliable operation. An exact mathematical definition is given of each of these phenomena; that they do actually take place is shown, qualitatively but rigorously, from the form of both sets of equations; methods for the quantitative solution of these are in part already available and will be discussed in a future work. Likewise, although we are actively engaged also in the study of the concrete aspects of the question, we shall limit the present report to an outline of the mathematical theory.

### C. NORMAL PHYSIOLOGICAL MODEL

Finally, we wish to emphasize that our machine does not purport to realize necessarily *an anatomical model* of the brain, that is, there need be no one-to-one correspondence between the anatomical neuron and the basic unit of the machine; we are concerned here only with the description of a *physiological model*, in which, as a whole, it is irrelevant whether the functions of a single neuron are taken up by a single unit or by a group of units in the machine, or vice-versa. Likewise, one could reproduce the functions of a circuit containing electronic tubes of various descriptions in terms, say, of a model circuit containing only triodes. We wish to emphasize also that our model intends to simulate the physiology of neurons in their *normal* condition in the living tissue, and not at all the various reactions they exhibit when tortured in the physiologist's laboratory: most of the latter will be as irrelevant to the study of the collective behavior of neuronal assemblies, as is the detailed knowledge of the radiation spectra of Na and Cl ions to the determination of the crystalline structure of the NaCl salt.

## 2. Symbols

$1(x)$ = unit step function
$\Sigma$ = Stieltjes integral or summation
$h, k, i, r$ = indices denoting integers (subscripts of superscripts)
$R, N$ = fixed integers
$a_{hk}^{(r)}, A_{hk}^{(r)}, b_{hk}$ = real numbers (coupling coefficients)
$s_h$ = real numbers (thresholds)
$t$ = time variable
$\tau$ = a fixed "time quantum"
$u_h(t)$ = piece-wise-constant functions (-0 or 1 in any quantal interval of time)
$\nu$ = class of all functions $u_h(t)$
$c_i(t)$ = a constellation of neurons at time t
$n_i(t)$ = number of neurons of $c_i(t)$
$E$ = a fixed set of neurons
$M_E, \overline{M}_E$ = classes of solutions of eqs. (2) relative to E.
$G(t)$ = group of transformations under which eqs. (2) are invariant
$\mathcal{N}$ = configuration space
$P(t)$ = representative point of system in $\mathcal{N}$
$\mathcal{F}$ = a functional space built over $\mathcal{N}$
$S(t)$ = a frame in $\mathcal{F}$
$\Theta_i, \Theta_{ij}, \ldots$ = patterns presented to, or constructed by, the machine

$f(a_{hk}; \lambda)$ = secular equation with variable $\lambda$
$\bar{\lambda}_h$ = eigenvalue of $f(a_{hk}; \lambda) = 0$
$\rho_{hk}$ = small random variation of $a_{hk}$
$\rho$, $<\delta\bar{\lambda}_h>$ = average values of $\rho_{hk}$ and of corresponding variation of $\bar{\lambda}_h$

## 3. Neuronic and Mnemonic Equations

### A. GENERAL REMARKS

1. The present considerations aim at simplicity, rather than at formal elegance; many restrictive assumptions are therefore made which could easily be relaxed, gaining thereby a greater apparent generality in our equations but, in reality, only complication which is better avoided at this early stage. The most evident is the fact that we use throughout summations instead of integrations, although Stieltjes integrals would be in many cases more appropriate to a faithful description of the anatomical situations of interest.

Instead of considering the actual speed of propagation of the neuronic discharge along dendrites and axones, we neglect the first and lump the second together with the synaptic delay into a single time-unit $\tau$; this is a better approximation than it may seem, because in the brain, as is well known, speed in axones is proportional to diameter and, although less generally, diameter to length. We schematize this situation by assuming that a neuron which receives a pulse (either does not fire or) fires after exactly $\tau$ sec; or, more generally, that $\tau$ denotes some conveniently small "time quantum", of which the neuronic delay times are (not necessarily equal) multiples (our neuronic equations (2), although apparently designed to describe only the first situation, also cover the second).

2. We shall base our treatment on two sets of equations: the *neuronic equations* (N.E.) which have constant coefficients and determine the instantaneous behavior of the system, and the *mnemonic equations* (M.E.) which account for the semi-permanent or permanent changes in the structure of the system caused by its past operation. This is, again, an artificial simplification of the actual situation, which is better described by retaining only the first set of equations, with coefficients taken as "slow" functions of time and past neuronic activity. The approximation thus made is analogous to the Born-Oppenheimer approximation of molecular physics, which consists in studying first the motion of the (much faster) electrons as if the nuclei were fixed, and then the behavior of the latter. It is justified physiologically by the experimental observation that electroshock, or concussion, cancels all memories of things learnt within a previous time interval of minutes or more, while memories acquired before

that time remain unimpaired: this makes it reasonable to assume that the brain takes about that much time to change the dynamical phenomena which we consider here to be the carriers of functional, short-range memories, into semi-permanent or permanent alterations. That the latter actually exist is proved by the fact that they are not suppressed by hybernation or artificially provoked cessation of all neural activity.

We may call this the *adiabatic learning hypothesis* (A.L.H.): the degree of adiabaticity of learning in the brain can be estimated roughly from the remark just made, with the conclusion that the engramming of permanent or semi-permanent memories takes roughly a time of $10^4$ to $10^5$ sec or more. The determination of the duration of semi-permanent memories in the brain is a task for experimental psychology, and is not discussed here.

The mathematical advantages of uncoupling the actual equations of neural activity into two distinct sets by means of the A.L.H. will be evident: by considering all constants frozen, the resulting N.E. are solvable notwithstanding their utter non-linearity, and in any case their very form leads immediately to many interesting qualitative conclusions, as we shall show later.

3. It is perhaps relevant to emphasize that the equations which we shall take as the basis of our treatment do not certainly contain, in themselves, any striking novelties. They are about what any neurophysiologist would write at once, should he wish to arithmetize, say, the kind of logic that is usually associated with neuronic circuits, or to formulate some reasonable guess about the growth of memory.

What we consider to be the essential point in our whole theory is, rather, the fact that arithmetization is considered here as the *necessary first step*: once equations are written, then, and only then in our opinion, the real groundwork can begin. Furthermore, equations alone mean very little to a mathematician; the detailed prescription of the type of information which is wanted from the solutions of a given equation constitutes a "problem", the formulation and solution of which is, in all cases, the most relevant question. We shall therefore be concerned here essentially with the formulation of problems which arise from these equations and are central to our theory of thought-processes; in so doing, we shall meet interesting and novel mathematical situations, the quantitative study of which is well under way and will be reported in the future. The qualitative discussions of Sections 4 and 5 will suffice for our present purposes.

## B. NEURONIC EQUATIONS

1. We take as the basic component of the machine—which for convenience we call a "neuron", although its functional relation to living

neurons need not be 1 : 1—a discriminator with a large number of inputs (dendrites) and a large number of outputs (branching axones). Signals can only travel *unidirectionally*, with infinite speed, from the output of a neuron to the input of the neurons connected to it; when a signal reaches a neuron it is annihilated, unless enough signals arrive with it to cause the neuron to fire a pulse, after a delay $\tau$, simultaneously in *all* its outputs. The intensity of these pulses may vary with the "anatomy" of the neuron, i.e. number of inputs, outputs, location in the machine, etc.; such pulses may be attenuated during propagation, or other phenomena may occur, as is discussed later. As a matter of formal convenience, we normalize all pulses to unit strength and account for larger or smaller strengths by giving suitable values to the coupling coefficients. Finally, a neuron will fire only if the total sum of afferent pulses is greater than its threshold. All coupling coefficients and thresholds are considered to be constant (adiabatic learning approximation).

We define the function:

$$1(x) = \begin{cases} +1 \text{ for } x > 0 \\ 0 \text{ for } x \leq 0 \end{cases} \tag{1}$$

let $u_h(t)$ denote a function belonging to the class $U$ of piece-wise-constant functions which are either constantly 0 or constantly 1 in any of the intervals $l\tau$, $(l+1)\tau$ ($l$ integer $\gtrless 0$); we take then as fundamental equations for the description of the instantaneous behavior of our machine (neuronic equations, N.E.):

$$u_h(t+\tau) = 1[\sum_{k,r} a_{hk}^{(r)} u_k(t - r\tau) - s_h] \tag{2}$$

The meaning of the coefficients $a_{hk}^{(r)}$ and $s_h$ is stated below; the anatomy of the machine at a given instant is described entirely by their values. (Taking $\rho_h\tau$ instead of $\tau$ at l.h.s. of (2) would not change the structure of these equations: an obvious re-naming of their coefficients would lead back to the form (2).)

2. $s_h$, usually $> 0$, is the threshold of the neuron $h$; the neuron $h$ fires at time $t+\tau$ if its *excitation* at time $t$ (given by the sum in (2) ) is greater than $s_h$.

$a_{hk}^{(0)}$ $(k \neq h)$ is the *coupling coefficient* that transfers the pulse originating from neuron $k$ to neuron $h$; it contains the *total effect* of the first on the second, *regardless* of the number of synapses between $k$ and $h$ and of the intensity with which the stimulus coming from $k$ reaches $h$ along each pathway. When $a_{hk}^{0} \neq 0$, we say that there is a (unidirectional) *direct channel* between neuron $k$ and neuron $h$, which causes a *facilitation* $k \to h$ if $a_{hk}^{(0)} > 0$, or an inhibition $k \to h$ if $a_{h}^{(0)} < 0$.

The rôle of the coefficients $a_{hk}^{(r)}$ $(h \neq k)$ and $a_{hh}^{(r)}$ is quite different:

$a_{hk}^{(r)}$ ($h \neq k$; $r$ integer $> 0$) is $\neq 0$ only if it is required that the actual mechanism of stimulation be such that the effect of the pulse from $k$ may reach $h$, or last on $h$ some time $> \tau$ after $k$ has ceased firing; this would be the case if stimulation were due, say, to some transmitter substance released at the synaptic junction, which would be re-absorbed only after a time $> \tau$. Such a mechanism would account for latency and be related to the well-known dependence of pulse frequency on intensity of stimuli. It may not be a bad approximation, in a model, to take $a_{hk}^{(r)} = 0$ for $h \neq k$, $r > 0$, except perhaps for input elements.

The coefficients $a_{hh}^{r}$ express instead the memory that the neuron $h$ retains of each of its firings (in the brain, for about $100\tau$ sec). For all we know, the characteristic observed shape of the neuronic discharge (as well as many other things) may well be only the result of biological necessity, and to ask that it be closely reproduced in a thinking machine might prove as binding as demanding that moving objects be built with legs rather than wheels. We shall want in any case $a_{hh}^{(r)} \ll 0$ for all values of $r$ from $r = 0$ until $r\tau$ becomes greater than the absolute refractory time of the neuron; for the latter and higher values of $r$ it may be convenient to follow different prescriptions, according as one wishes to study the actual behavior of the brain on this model, or instead to construct a thinking machine for some special purpose.

3. As an example (among the many that might be produced) of the fact mentioned earlier that our N.E. might be a poor description of the anatomy and yet give a faithful description of the physiology of a nervous system, we consider here the situation that would arise if, in a nerve, or bundle of fibers, the electrotonus due to axones which are carriers of pulses should induce firings in other axones of the same nerve which originate from neurons that *have not fired.*

This possibility was not contemplated when writing the N.E. (2). A model which reproduces also this new type of behavior must lead to equations such that signals can be either *transmitted directly* from neuron $k$ to neuron $h$, or *induced* into the channel $k \to h$ by the firing of some neighbouring neurons; the neuron $h$ must not be able to discriminate whether the pulse it receives through that channel has a direct or induced nature. Taking for simplicity $a_{hk}^{(r)} = 0$ for $r > 0$ ($h \neq k$), we obtain clearly the wanted equations by replacing $\sum_r a_{hk}^{(r)} u_k(t - r\tau)$ in (2) with

$$a_{hk}^{(0)} u_k(t) + \sum_{k_i \neq k} b_{hk_i}^{(0)} u_{k_i}(t) \tag{3}$$

where $\sum_{k_i \neq k}$ means sum over the neurons $k_i$, neighbours of $k$, the axones of which can act in this way on the channel $k \to h$, and $b_{hk_i}^{(0)}$ are some suitable coefficients.

It is then evident that, renaming the coefficients, one finds again N.E of type (2). The same can be said for inter-dendritic interference.

### C. MNEMONIC EQUATIONS

1. There is sufficient evidence to prove that memory in the brain is due both to functional processes and to reversible and irreversible alterations of its micro-structure. Very little, if anything, is known for certain beyond this, so that we are forced to rely upon "plausible" hypotheses if we wish to assign the specific laws which determine semi-permanent or permanent physico-chemical changes. We shall not hesitate to do so for the sake of concreteness; we wish however to emphasize that the qualitative analysis of thought-processes which is the purpose of this work does not require precise knowledge of these laws, but only that they share some very general features, which may be assumed with much greater reliability.

Here lies a substantial difference between the brain and the thinking machine: the latter, which is obviously not restricted by the severe limitations of biological necessity, may have mnemonic devices and laws much more efficient than those of Nature, while giving rise to thought-processes (as described by the N.E.) of the same type. We feel also that, as regards memory growth and contrary to the situation that arises in the study of the N.E., a thinking machine of this sort might be of greater use to neurophysiology than vice versa; observations performed on models which can be built with mnemonic laws changeable at will, might help to shed light on the quantitative aspects of biological phenomena which are extremely difficult to observe directly.

Thought-processes in a portion of the cortex may be ascribed either to excitation of neurons which would be otherwise mostly at rest, or to inhibition of the activity of neurons which would be otherwise unceasingly firing. To the first one would associate mnemonic mechanisms which make firing easier with the progress of learning (this we may call a *facilitatory*, or *positive*, type of memory); the opposite with the second (*inhibitory*, or *negative memory*). Both types offer interesting possibilities for machine construction; since they obey essentially the same kind of N.E., we refer here throughout only to the first type.

The so-called "genetic", or "anatomical", i.e. permanent inherited memory, corresponds clearly in our description to the fact that, as we shall see, some (actually most) of the coefficients which couple neurons together must be taken initially, and kept throughout, vanishing. Our mnemonic laws will therefore be chosen so that if a coupling coefficient vanishes initially, it stays forever so, while its modulus may grow to maximum value from any given initial non-vanishing value.

In our model, thought-processes will be represented by non-trivia

solutions of the N.E.; the machine can also "think" therefore if all coefficients in the N.E. stay forever frozen, i.e. if the machine cannot learn or forget, provided these coefficients have convenient values. The present framework can thus account, as it should, for a clear distinction between "instinctive" and "intelligent" behavior. It is natural to suppose that genetic patterns determine the laws according to which cells duplicate, branch out and anastomize, rather than the actual ultimate detailed anatomy of a tissue (thus, a "gene" carrying the instruction "add + 1" would suffice to generate all integers from zero, while an infinite number of "genes" would be obviously required if each integer should have its distinctive "gene"); then even a few mutations may determine the appearance of neural structures quite at variance with previous patterns, from which the evolutionary laws can secure the selection of the fittest, that is those which possess the most favorable neuronal couplings. Our definition of thought comprises thus two types of performance for which we use the conventional terminology: "instinct", which is learnt genetically, and "intelligence" which arises when these couplings can change during the life of the individual.

2. The quantities $s_h$, $a_{hh}^{(r)}$ and $a_{hk}^{(r)}$ $(h \neq k)$ were seen to play quite different rôles. When assigning their variation with time, we refer henceforth to a machine rather than to the living brain, for the reasons mentioned before.

It is apparent from (1) and (2) that the maximum learning capacity of the machine is already reached by assigning suitable variations only to $a_{hh}^{(r)}$ and $a_{hk}^{(r)}$. Once the mnemonic laws are given for these, changes induced in the $s_h$ appear as the best way of controlling the operation of the machine. We shall return on this point in Section 4 and consider here the $s_h$ as quantities the values of which do not change because of mnemonic laws, but, if at all, through some different mechanism.

The coefficients $a_{hh}^{(r)}$ have already been discussed in B, 2, p. 210; for the purposes of the present discussion we may assume.

$$a_{hh}^{(r)} = \begin{cases} -\infty & 0 \leq r \leq R \\ 0 & r > R \text{ (integer)} \end{cases} \tag{4}$$

For $h \neq k$ a convenient law is (for positive, or facilitatory $a_{hk}^{(r)}$):

$$\frac{da_{hk}^{(r)}(t)}{dt} = \{\alpha^{(r)} u_k(t-\tau) u_h(t) - \beta^{(r)} \mathrm{I}[a_{hk}^{(r)}(t) - a_{hk}^{(r)}(0)]\} a_{hk}^{(r)}(t) \mathrm{I}[A_{hk}^{(r)} - a_{hk}^{(r)}(t)], \tag{5}$$

where $\alpha^{(r)} \gg \beta^{(r)} > 0$, $A_{hk}^{(r)} > 0$, and it is imposed that $a_{hk}^{(r)}(t)$ be continuous, with $a_{hk}^{(r)}(0) \leq A_{hk}^{(r)}$.

For the sake of concreteness we take (5) as the mnemonic equations (M.E.) of our machine; we also neglect inhibitory (negative) couplings, to which (5) is immediately extended in an obvious manner. We may suppose here, for simplicity, that only coefficients with $r = 0$ survive, and that all $A_{hk}^{(0)} = A$ and all $a_{hk}^{(0)} = a$. We have already emphasized that all that we actually need are M.E. that admit solutions having the same qualitative behavior as those of (5); these we proceed to discuss briefly.

3. We write $a_{hk}(t)$ for $a_{hk}^{(0)}(t)$. The M.E. (5) describe a situation in which $a_{hk}(t)$ never becomes smaller than $a > 0$, nor greater than $A$. When the latter value is reached, it is retained for ever: the information is engrammed permanently. This is perhaps an oversimplified view of the real situation in the brain; it could be, though, easily modified.

$a_{hk}(t)$ increases if, and only if, the neuron $h$, which is connected by a direct channel to neuron $k$, fires at time $t + t$ and has received a pulse at time $t - \tau$ from the latter. It decreases slowly afterwards ($\alpha \gg \beta$), until the same situation repeats. Only if a series of such rises occurs, without excessive delays in between, can $a_{hk}(t)$ reach the engramming value $A$.

There is ample choice of mechanical devices which can reproduce qualitatively this behavior. If it is desired that the machine exhibit a behavior typified by (3), coefficients like $b_{hk}^{(0)}$ might be given constant values, not subject to mnemonic phenomena.

## 4. Qualitative Discussion

### A. Operational definition of "thought"

1. We propose now to show that, as was mentioned in the Introduction, a machine that works according to the N.E. (2) and the M.E. (5) will exhibit phenomena which are typical of a nervous system, provided of course the number of its elements is sufficiently large and the initial values $a_{hk}^{(r)}(0)$ of the couplings among these ("genetic memory") are conveniently chosen (e.g. so as to prevent "epilepsy": cf. C, 4, p. 221).

The most obvious features of the N.E. are non-linearity and unidirectionality of pulse transmission; their solutions describe therefore in any case states of excitation (or "motions", or "modes") that "travel" unidirectionally from neuron to neuron and interfere nonlinearly whenever they meet. This interference is either instantaneous or nearly so, as it happens when summation of pulses at the synapsis of a neuron causes its firing (as described in the r.h.s. of (2)); or delayed, as it happens when pulses, which would otherwise cause the firing of a neuron, cannot do so because they reach that neuron when it is still inhibited by a previous firing, due to different pulses.

2. *We define a thought-process*, operationally, *as a solution of the N.E.*, or, equivalently, as the corresponding "motion" in the machine. It is convenient further to qualify this definition, so as to meet obvious objections.

We may disregard as "trivial" and not consider as "thoughts" solutions that correspond to (say accidental) firings of neurons at a given time, such that no other neurons are induced into firing thereby and all activity ceases immediately afterwards. Any "thought" implies thus the passing of at least one neuronic channel.

3. For any given set $E$ of neurons, all motions of a given duration can be classed either into a set $M_E$, the motions of which cause at least one neuron in $E$ to fire at least once, or into a set $\overline{M}_E$ of the remaining possible motions. There is thus (and in many ways) the possibility of establishing operational distinctions between "types of thought"; should, for instance, a portion of the machine correspond to the central and one to the autonomic nervous system, the name "thought" could be further restricted thereby to the solutions of the N.E. which affect only the neurons of the first. If, in a different partition, $E$ is the set of neurons the firing of which is associated somehow with consciousness (e.g. because they control a loudspeaker, or some prescribed feed-back mechanism), then all motions of $M_E$ can be taken as representing the "conscious activity", all those of $\overline{M}_E$ the "subconscious activity" of the brain.

It is interesting to remark that, in the latter instance, because of the various possibilities of interference discussed before between the motions of $\overline{M}_E$ and those of $M_E$, each type of activity influences the other. "Psychoanalysis" reduces for this machine to a simple and well-defined mathematical problem.

### B. PROBLEMS CONNECTED WITH THE N.E.

1. The N.E. clearly contain, as special cases, the description of all logical networks of the kind beautifully analyzed in the pioneering work of McCulloch and Pitts. Should their solution be attempted by the obvious method of iteration, they would, for these cases, give just as much—or as little—information as can be gathered from the standard logical switchboard analysis; there is here a clear analogy with the Darboux (better than the Cauchy) problem of the theory of differential equations.

The systematic algebraization of logic, which is the real content of the N.E. (with frozen coefficients), permits us to pose for them much more general questions, which may be treated with a variety of mathematical tools; the logic of the system is seen to play a rôle so to speak similar to that of the constraints which limit position and mobility of a dynamical system; an appropriate treatment of the N.E. will permit, as with the

equations of motion of dynamical systems, the search for those long-range collective solutions which, in our scheme, form the basis for a useful analysis of thought-processes.

2. We consider first of all the N.E. with frozen coefficients, in keeping with the A.L.H. Their quantitative discussion poses some interesting and novel mathematical questions, and will probably require the introduction of techniques *ad hoc*; on the other hand, it is evident that in simple cases, such as may correspond to situations involving very few neurons, the N.E. may be solved on inspection. It is also clear that straightforward combinatorics can give useful information on the possible types and multiplicities of the solutions of interest, as defined below; and that this can be translated at once into the customary language of "excitation probabilities", etc. While deferring to future reports for detailed studies on these matters, on which work is in progress, it is fully sufficient for our present purposes to formulate the "problems" which we envisage as most relevant in study of the N.E., and to discuss them briefly at a qualitative level.

The first obvious, and obviously important, remark is that the N.E. are not uniquely determined; their formulation (2) is perhaps deceptively simple. Because of the definition of the function $\mathrm{I}(x)$, there is a whole group $G$ of transformations which change a given set of N.E. into an *equivalent* one—having, that is, exactly the same solutions, although not necessarily the same form (thus, $\mathrm{I}(x) = \mathrm{I}(2x) = \mathrm{I}(x^3) = \mathrm{I}(\sin x)$, etc.). This fact was already used in the discussion of the threshold values $s_h$ made in Section **3,** C, 2; it shows, for instance, that matrix algebra should be used with caution in handling these questions.

For the same reason, "suitably small" changes of the $a_{hk}^{(r)}$ and $s_h$ will not change the solutions of a set of N.E.: this adds credit to the reliability of the A.L.H. and provides what we may call the *first criterion of stability* of the machine.

3. We shall soon specify what types of "input" and "output" seem most appropriate for a machine of this sort; we are now interested in the "spontaneous" activity of the machine, which we define as that which takes place in it when, at a given time $t_0$, the machine starts from any given state of excitation and no input pulses are fed into it for $t > t_0$.

In a linear network—it is convenient, for purposes of comparison, to refer to a system of harmonic oscillators with linear couplings—such activity is naturally analyzed in terms of eigensolutions, eigenfrequencies, harmonics; the behavior of a single element is in general not periodic, but simple periodal analysis will resolve it into a sum of periodic normal modes, which have a collective character and may be defined as the motions of quasi-particles (this remark already suffices to eliminate as

illusory any attempt at deciding on the existence of periodic motions in the brain through observations performed on one or few neurons). In a non-linear system things become much more involved; e.g. one finds in general, besides harmonics (multiples of a fundamental frequency), also subharmonics (multiples of a fundamental period).

The extreme schematization which is expressed by the form (2) given here to the N.E. has the evident consequence that one can only expect subharmonics; if there are periodic solutions of the N.E., these are *reverberations*, i.e. transfers of excitation from neuron to neuron which may reach anywhere into the machine and, after the closing of suitable multichannel paths, repeat with a periodicity which is, obviously, some integral multiple of $\tau$.

The consideration of reverberations is central to our approach. There are tremendous numbers of them even in the simplest conceivable models; their types, paths, multiplicities are determined by the coefficients $a_{hk}^{(r)}$ and $s_h$ of the N.E., and change therefore, because of the M.E., with learning and forgetting.

4. The first mathematical problem is therefore the determination of all the solutions that correspond to reverberations, or free ("spontaneous", "autogenic") modes compatible with the N.E. The minimum duration of a reverberation is clearly determined by the refractory period of the neurons through which it travels; if we assume that "normal" activity (i.e. without special stimulation) of the neurons in the brain uses the total period of the pattern of spike-afterpotentials ($\sim$ 100 msec), then the maximum possible frequency (reverberations involving $\sim$ 100 neurons) is about 10 cycles/sec, which coincides with the frequency of the $\alpha$-waves of the E.E.G. If we assume further that stimulation may force the neurons of the brain into using a refractory time intermediate between the absolute ($\sim$4 msec) and the total time, then this maximum frequency increases and the number of neurons necessary for the smallest permissible reverberations decreases. We do not wish to draw any conclusions at this early stage from these remarks, which may be a gross oversimplification of reality; we only state here that they are not in disagreement with observation.

If thresholds and couplings have the values that are observed in the brain, then reverberations certainly involve several tens of neurons. Reverberations, furthermore, should last for ever in an ideal machine, a conveniently long time in a real machine and in the brain. From the first remark it follows that one cannot expect to observe direct evidence of prolonged autogenic activity in a portion of the cortex in ordinary conditions: this would require innumerable microelectrodes stuck into as many neurons for an experiment to be feasible. One would expect, how-

ever, from our theory, that if thresholds are sufficiently lowered artificially or the intensity of stimuli increased, then also a very small number of aptly chosen neurons should suffice for a prolonged autogenic reverberation to take place. In a brilliant series of experiments A. Féssard (Symposium on Memory, Naples, 1960) has demonstrated, by using tetanic potentiation, that this actually happens: he recorded reverberations among only four neurons which would last minutes. We regard his results as a crucial, if only partial, confirmation of our theory, which was developed while we were still unaware of his work.

5. Reverberations, as all other motions, interfere non-linearly unless one reverberation never affects in any way the neurons of another, i.e. as we shall say, is *disjoint* from the other. At this point the analogy with a linear network breaks down completely, much to the advantage of our machine, which possesses many more essentially distinct modes of behavior than a linear system. It is still possible to classify *all* possible spontaneous reverberations, for instance according to periodicity, multiplicity (i.e. degeneracy), etc.

The next mathematical problem that arises is the study of the evolution of the state of excitation which was present in the machine at time $t = t_0$, as was discussed in 3, above. It may either coincide with a configuration of excitations which characterizes at $t_0$ a reverberation, and thereafter continues its periodic behavior; or, more often, *decay*, into a reverberation, or *develop* into a reverberation, or produce *catastrophic behavior*, i.e. lead to total (or nearly total) simultaneous excitation ("epilepsy") of the neurons, which may decay immediately afterwards into rest (cf. the N.E.).

Excluding for the time being the last dramatic alternative, we find here the most interesting situation, as close an analogue as is possible with a non-linear system to harmonic analysis. In a frozen state of knowledge out of the machine as many distinct responses can be evoked as there are distinct excitable reverberations, or modes; each of these we may identify with a "pattern" which the machine knows genetically, or has learnt; the "initial configuration" is the pattern which is presented to the machine the set of (one or more) disjoint reverberations to which the latter gives rise (depending upon the value of the couplings) is the *analysis of that pattern* performed by the machine which corresponds to the state of knowledge it has learnt until that moment.

Apart from learning, we have here the counterpart to what we regard as the essential activity of the mind, the ability to analyze a situation, or shape, or pattern, into a set of already classified patterns. No single element acts as a classifier since the total response of the machine is required for this analysis.

6. The situation described above is manifestly an extreme simplification. The next mathematical problem is in fact that of studying the evolution in time of the total state of excitation of the machine when its "input" is subject to continued external stimulations.

It will also be expedient, of course, whenever dealing with very large assemblies of neurons, to distinguish between "traveling" and "stationary" solutions. We have been considering thus far only the latter, but it is clear that, as soon as distinctive special-purpose "organs" are built into the anatomical structure of the machine, our previous considerations should be restricted mostly to the latter, with pulses travelling from organ to organ as among the boxes of a diagram.

It is also to be expected that there will be a maximum duration, and a maximum complexity, beyond which reverberations cease to be significant for pattern-analysis. This assumption, or requirement, will greatly facilitate the mathematical study of the problem formulated in this section.

7. We have thus far taken, for the sake of simplicity, a perhaps too realistic view of reverberations as modes which are actually connected to fixed chains of neurons. This is not certainly the case when one considers the normal modes of linear networks, and it is therefore of interest also to investigate the possibility of resolving actual motions, which do not have manifest periodicity, into "normal" periodic collective modes (cf. 3, above: the Lissajous figures of linear problems are an example of this behavior); any such latent periodicity would be easily revealed by observations made upon populations of neurons (e.g. with the E.E.G.).

Questions of this nature, and many others, suggest themselves in a quantitative investigation; they need not be considered here in further detail.

### C. RÔLE OF THE M.E.

1. In the preceding section we have focused our attention on the operation of the machine when all coupling constants and thresholds are kept fixed, and have found that *reverberations* play a central rôle in its most typical activity, which is *pattern-analysis* in a very general sense. All such statements presuppose already, of course, the existence of favorable conditions, as are expressed for instance by the assumption (A, 1), which prevent epileptic, or catastrophic, behavior; it was also implicitly assumed that the machine is indifferent to the "meaning" (referred to any standards) of what it knows genetically or has learnt during its past activity. While we can reasonably expect that careful engineering and a long series of painstaking adjustments would in the end produce devices capable of some useful performance solely by virtue of conveniently chosen N.E., we are

much more interested in machines that can adjust themselves to prescribed tasks by means of some learning mechanism; this should also give the machine a tendency to organize itself into increasingly reliable operation, so as to compensate for minor flaws in the accuracy of its elements.

The M.E. provide, to a large extent, the answer to these questions, as we shall now show. In the course of the same discussion it will become apparent, however, that a machine of this sort is not realistically conceivable unless at least two additional controlling devices are not also explicitly included; the first we identify tentatively with the thalamus, the second, with more assurance, with the reticular system of the brain. The necessity of devices of this sort, if not already suggested in the brain by anatomical and physiological evidence, is made imperative in the machine by the structure of the N.E. and M.E.

Mentioning a "thalamus" takes us, of course, one step nearer to the "sentient" machine than we wish to stay for the time being; we shall therefore restrict this part of our discussion to barest essentials, pointing only to what is relevant for purely "rational" thought.

2. There are many mathematical ways of representing the overall situation and evolution of the machine, each suited to some special purposes. We mention here briefly a few which take the instantaneous state of each neuron as the object of interest.

If the number of neurons in the N.E. is $N$, then a solution of the N.E. at time $t$ is representable by means of a one-column matrix with $N$ rows, the element of row $h$ being given by $u_h(t)$; or one can define, equivalently, an $N$-dimentional *configuration* (or *neuron*) *space* $\mathcal{N}$, which has $N$ axes, on the $h$th of which the abscissa is $u_h(t)$. The state at time $t$ of the machine is thus represented by the *point*, or *matrix*, or *vector*, $P(t) \equiv \{u_h(t)\} \equiv \vec{u}(t)$, its evolution in time by the (discontinuous) motion of the point $P(t)$.

All trajectories in $\mathcal{N}$ are invariant under the transformations of the N.E. which belong to the group $G$ defined in B, 2, provided of course that at each time $t$ one takes $G$ as it is determined by the M.E.: now, $G = G(t)$.

$\mathcal{N}$ contains at most $2^N$ points; a trajectory in $\mathcal{N}$ is a polygonal joining some, or all, of these points. A reverberation is represented in $\mathcal{N}$ by a *closed polygonal* (and lasts at least as long as the coefficients in the N.E. stay frozen).

$P(t)$ changes in $\mathcal{N}$ (i) because it describes the evolution in time of a solution of the N.E., (ii) because the N.E. themselves change, due to the intervention of the M.E. The A.L.H. permits the study of simple phenomena by separating step (i) from step (ii): it allows, that is, that they be performed alternately. Step (ii) becomes necessary as soon as, because of M.E., the N.E. undergo a transformation which does not belong to G.

A qualitative discussion is better stopped here (see, however, Section 5, A, 2); it should already be clear from what little has been said on this subject, though, that the introduction of spaces of functionals on $\mathcal{N}$ will be of the highest conceptual importance, because then everything becomes again linear, group theory may be resorted to as a valid tool of analysis, each "pattern" is easily made to correspond to a point, and problems such as those of language translation or study of emotive behavior can receive a precise mathematical formulation.

3. The pattern-analysis described in B, 5, presupposes, clearly, that the machine has already formed, either genetically or by learning, some typical responses (or modes, or patterns), in terms of which a pattern presented to it is analyzed. Very little, if anything at all, can be expected from a machine with fixed constants built entirely at random: the most likely thing to occur in such a case is that, unless the experimenter arranges the connections of the machine in a way that is equivalent to giving it a genetic memory, the only resulting effect will be a total loss of information. Also a machine endowed with ability to learn will give, at best, a poor performance, unless the controlling devices mentioned above (1) are included into it, as we shall soon discuss. The best procedure, or at least by far the most economical, appears to be in any case that of borrowing as much as possible from anatomical and physiological information.

Before proceeding further we need say a few words about the kind of "input" and "output" which seems appropriate to a machine of this nature. The notions of input as "that which comes before" and of output as "that which comes after" the machine proper are clearly out of the question; we are interested in what occurs *at all places* and *at all times* in the machine, and a relatively small number of terminal plugs could never tell us readily this much. Adequate inputs and outputs are instead devices out of which (inputs) an afferent lead goes to *each* neuron of the machine, or into which (outputs) an efferent lead comes out of *each* neuron; or, more economically, these leads connect input and output terminals with a *large* number of neurons spread throughout the machine, and connected with all the other neurons so that no relevant information on its behavior is lost. Anatomically, this seems to correspond to some regions of the brain stem for the afferents, for instance, to the thalamus. It is clear that such a device is what can be best desired for many sorts of feed-back operations; to this point we shall return briefly in connection with the learning mechanism.

4. We have excluded, in the discussion made in B, 5, the possibility that a stimulus presented at the input may produce catastrophic, or epileptic behavior. In any real machine constructed with a very large number of elements and connections among these, however, and even with very good

engineering and planning, catastrophic behavior is the very first thing to be expected.

Even assuming ideal starting conditions, the intervention of the M.E. will soon change the values of the $a_{hk}^{(r)}(t)$. The maintenance of a reverberation presupposes that, after its cycle is completed, the same situation repeats identically. Unless the ratios of the numbers and weights (as given by the $a_{hk}^{(r)}(t)$ ) of output $v$. input terminals are kept within very critical limits (cf. the operation of a nuclear reactor), the most likely event to occur is that the initial situation does not repeat exactly after one cycle of reverberation is completed, but more, or fewer neurons are excited than the correct number. In either case a process very similar to a chain reaction might take place immediately, that is, a very fast excitation of all neurons, or a very fast extinction of all activity; more generally, totally uncontrollable phenomena would occur as a rule rather than as an exception.

Whenever learning is involved, and in any case whenever design is not ideally perfect, a controlling mechanism (like the cadmium bars in reactors) is necessary to prevent any such possibility. We saw in Section **3,** C, 2, that we can use, without any loss in the learning capacity of the machine, the neuronic thresholds $s_h$ for this purpose; this is also, clearly, the best choice from a practical point of view.

Any machine that works according to the N.E. and the M.E. necessitates therefore a mechanism which, upon receiving information on its local and general activity at a given time, may alter the thresholds $s_h$ of the neurons, so as to avoid catastrophic conditions at later times. We have described here the function of the reticular system of the brain, as was made clear by the profound physiological investigations of G. Moruzzi & H. W. Magoun. The existence of such a mechanism will provide the *second criterion of stability* for our machine.

It will also be necessary, of course, to resort to the methods of which Nature avails herself in the brain: most of the $a_{hk}^{(r)}(0)$ will be taken as vanishing, in such a way that a neuron be connected through its efferent terminals mostly to rather distant neurons. Thus, excitations spread out and tend to stay below the epileptic thresholds; if couplings were only with "nearest neighbours", epileptic waves would be the only mode of operation of this machine. Such a choice of initial coupling coefficients will provide the *third criterion of stability*. We may include into this the action of inhibitory couplings also.

5. We can now discuss the operation of the machine from the point of view of the M.E.; learning and forgetting will play an equally important rôle. The cortex of the brain has many, more or less specialized input and output "areas"; we shall refer, generically, only to "input" and "output" and refrain from using the current terminology of "sensory", "motor"

and "associative" zones, which presupposes a more detailed structural knowledge than is needed for a qualitative discussion.

We also forgo such obvious things as the convenience that special input devices be constructed so as to perform a preliminary analysis of the figures, or patterns, "shown" them (fed into them; we refer, for concreteness, to a visual input); thus, if a set of homothetic triangles is shown, one may require that the input device transmit to the direct input of the machine only the image of a standardized triangle, plus an information on the value of the homothety parameter. Devices of this sort are not hard to conceive as special organs or extensions of the machine itself, to which they are linked by additional N.E. with fixed coefficients (learning is undesirable at this level). We refer hereafter only to the direct input of the machine, and assume that any such simplification has already been performed somehow.

Without a thalamus, the machine can only learn *by repetition* from habit. Suppose it starts as a *tabula rasa*, i.e. with all couplings having values $a_{hk}^{(r)}(0)$; suppose also, for the sake of simplicity, that a figure may be presented to it any number of times, but each time only for a very short duration. Each presentation will stimulate into firing a number of input neurons which we assume to be sufficiently large to initiate a collective activity, which spreads into the machine as described in A, 1 and B, 3 and 5. Unless the collective motion thus induced is strongly favoured by the genetic memory (i.e. the $a_{hk}^{(r)}(0)$; this ought to be the case for such things as the infant's sucking reflex), nothing much should happen after the virgin machine sees a figure $\Theta_1$ (say, a triangle) for the first time; reverberations will be evoked but, because of the low values of the coefficients $a_{hk}^{(r)}(0)$, a great many pulses are required to cause the firing of each single neuron, so that periodic, or nearly periodic, modes may be expected to involve a great many neurons and to be quite slow. The activity thus induced will be rather *diffuse*, not yet quite specific; it will last for some time $\gg \tau$ and the coefficients will start changing slowly because of the M.E. The pattern of this motion will be altered as soon as this change requires step (ii) as described in 2, p. 220: the increase of the coupling constants will tend to favor the creation of pathways through which reverberations are facilitated, i.e. have shorter cycles and require less neurons. Normal biological disturbances, or mechanical variations, spontaneous (e.g. due to the variations of the coefficients caused by the M.E.) or imposed, e.g. by the reticular system, will cause this state of motion eventually to cease; the chance that a single direct channel be used as many times as is required by the M.E. for a permanent engramming of a facilitation is very low; what alterations have been caused by the sight of $\Theta_1$ will be slowly forgotten by the machine.

Clearly, however, the thing is quite different if $\Theta_1$ is shown many times in succession to the machine: then the semi-permanent changes induced by the M.E. will accumulate, until permanent changes are induced that definitely facilitate the "most convenient" reverberations evoked by $\Theta_1$, i.e. those that are quickest and involve the least number of neurons; these we regard as specific to $\Theta_1$. Permanent engramming is favored if the intervals between the exposures of $\Theta_1$ are made shorter, disfavored otherwise; it may however take a very long time (days, or months, or years in an animal), or never occur at all: semi-permanent memories may decay very slowly indeed, and there may also be more stages, with various decay times, before permanent engramming obtains, than is assumed in the form (5) of the M.E.

It is evident that if a series of random figures is shown in succession to the machine, it will learn nothing, except perhaps a response meaning only that "a figure is being shown" (see later for a further discussion of this fact); but if among these random figures a given one $\Theta_1$ is shown repeatedly enough, the machine will "learn" only $\Theta_1$, together with the general response mentioned before.

This discussion is manifestly incomplete; diminishing response because of assuefaction, for instance, is not considered. The reason we do not wish to push it farther is that it appears too easy, rather than too difficult, to answer such questions at a qualitative level—which may be as dangerous as inconclusive. For instance, both the action of inhibitory couplings (which we have arbitrarily disregarded here for the sake of brevity, although their importance cannot be doubted) and of external inhibitions, e.g. from the reticular system may be invoked to explain assuefaction. Even with these restrictions, we hope that the present discussions of the subject may suffice to provide a convenient basis for further elaborations.

6. Learning by repetition, as described in 5 above, does not appear very satisfactory. One may say that the virgin machine has no more reason to wish to learn a triangle than the infant child. The same mechanism, however, makes it easy to explain "learning by punishment and reward", or *conditioning of the first kind* (as we shall say) and to account why the latter is much faster than the first.

The machine must, of course, be sentient to some degree to know whether it is being punished or rewarded. This will mean here that the thalamus has in-built criteria (homeostatic devices) which enable it to "like" or "dislike" what it records. Suppose that this is the case, and that the thalamus can either suppress, or create a state of excitation in the neurons. Then, even when a stimulus is presented only once at the input, the thalamus may evaluate the situation through its homeostats and determine either a quick suppression of it (or of a part of it, e.g. that

which produces some specific motor pulses), or that it be reinforced and maintained until permanent engramming is achieved, much more quickly in this way, clearly, than by "external" repetition at long or random intervals.

A thinking machine may perhaps do without this device and method of learning, which becomes of utmost importance only when survival has to be fought for.

7. We have discussed in 5 (p. 222) the situation that occurs if a given figure, or pattern, $\Theta_1$ is presented repeatedly at the input. We consider now what can happen if two distinct patterns $\Theta_1$ and $\Theta_2$ are used.

If the machine has already learned $\Theta_1$, i.e. if it responds to $\Theta_1$ with specific reverberations, when $\Theta_2$ is presented for the first time the situation is not the same as at the origin of time; some facilitations are already formed in the connections, so that the machine, as its first reaction, will show all of the reverberations which $\Theta_2$ may evoke in common with $\Theta_1$, plus extra motions which will slowly be changed into a response distinctive of $\Theta_2$. In other words, if the machine knows already $\Theta_1$, and sees $\Theta_2$ for the first time, its only immediate response will be to tell how much $\Theta_2$ has in common with $\Theta_1$; later on it will learn also $\Theta_2$. Thereafter, it will be able to analyze likewise $\Theta_3$ in terms of $\Theta_1$ and $\Theta_2$; and so on.

This is, at its simplest, the mechanism of pattern-analysis, as it develops with the evolution in time of the machine—its "education". The analysis the machine is capable of performing at time $t$ is determined, because of the A.L.H., solely by the N.E., taken with the values their coefficients have at that time.

8. The most typical and distinctive characteristic of the human mind is, in our opinion, its ability to abstract what is "common" to two, or more, situations or patterns, and to retain the result of this operation as a new pattern, which is entrusted to the memory as if learnt from the outside. Just as pattern-analysis was seen to be the fundamental operation of a machine described by N.E. with frozen coefficients, we proceed now to show that *abstraction* is the other fundamental operation performed by a machine which obeys N.E. and M.E. as well. The same discussion will clarify also the exact meaning to be given to the word "common" used above.

At any time $t$ intermediate between the time $t_0$ at which a pattern $\Theta_1$ is activated into the machine (e.g. a triangle $\Theta_1$ is shown at the input) and the time $t_1$ at which the collective motion aroused by that act dies out ($t_1 - t \gg \tau$; no other such activations are supposed to take place, and the mechanism described in 7 above is excluded), there will be $n_1(t - t_0)$ neurons which actually fire, and form the *constellation* (set) $c_1(t - t_0)$ engendered by the activation of $\Theta_1$ at $t_0$. There will be also, in addition, a

constellation $c_1'(t - t_0)$ of $n_1'(t - t_0)$ (in general, $n' > n$) of "penumbral" neurons which *do not fire*, but receive a *subliminal facilitation* from the activation of $\Theta_1$ at $t_0$.

Suppose now that at some time $t_0'(t_0 < t_0' < t_1)$ another pattern $\Theta_2$ is activated (a different triangle $\Theta_2$ is shown at the input). The machine may have already learnt $\Theta_1$ in a permanent or semi-permanent way, or not. Define likewise $n_2(t - t_0')$ and $c_2(t - t_0')$, etc., and neglect (this is incorrect and could easily be avoided in this discussion without altering its *qualitative* conclusions) the alterations caused by the non-linear interference of these motions at times between $t_0'$ and $t$ in the constellations $c_1'(t - t_0)$ and $c_2'(t - t_0')$; suppose that at least for some $t$ between $t_0'$ and $t_1$ the intersection of the sets $c_1'(t - t_0)$ and $c_2'(t - t_0')$ is not void and contains a constellation $c_{1,2}(t - t_0')$ of neurons which (would receive only subliminal stimuli from *either* motion, but) *fire* because the non-linear summation of the stimuli from *both* motions exceeds their thresholds: then, $c_{1,2}$ cannot be distinguished from a state of excitation such as would be produced by the presentation (not necessarily at the direct input which this, we recall, has the same structure as the rest of the machine) of a pattern $\Theta_{1,2}$. $\Theta_{1,2}$ can be aroused, clearly, only if $\Theta_1$ *and* $\Theta_2$ are shown at the input, the second after the first, and will result quite differently, in general, from $\Theta_{2,1}$, if the temporal development of both motions and the interference effects between them are treated without the illegitimately oversimplified assumptions which were made here for short.

Although presented here in mere outline on purpose, this line of reasoning makes it evident that, because of the non-linearity of the N.E., whenever a pattern is activated while the response to a previous one has not yet died out, the machine can *abstract* something which is "common" to both (the structure of the machine decides what meaning this word should have) and then adjust itself through the M.E. so as to memorize it, permanently or not, not differently from its normal behavior in response to any other pattern.

"Patterns of patterns" of any sort, in any number, may be formed and learnt in this way: chains of abstractions can take place without limitations other than those imposed by the complexity and structure of the machine. We recognize here, in full, the mechanical analogue to the faculty of abstraction of the human mind.

9. Abstraction alone is not enough: it would be highly uneconomical to remember all single instances, once the general concept is grasped. Our machine takes, in this respect, good care of itself. Suppose that, say, a sufficiently large number $a$ of random triangles $\Theta_i$ have been shown to the machine, which has learnt them semi-permanently; it has also formed the responses $\Theta_{1,2}$, $\Theta_{2,1}$, $\Theta_{1,2,3}$, ..., $\Theta_{1,2,3\ldots a} \equiv \Theta$ which are common to all sub-

sets of triangles (ordered or not) and memorized these semi-permanently. This memorization is accompanied by a process of facilitation; whereas, before memorization, $\Theta$ could be evoked only if *all* the triangles were shown to the machine, after facilitation, the $a_{hk}^{(r)}$ that appear in the N.E. which describe the behavior of the neurons of the constellation $c_{1,2,\ldots,a} \equiv c$, have increased their values considerably; suppose, for maximum simplicity, these values to have become so large that each single motion $\Theta_i$ gives now to the neurons of $c$, instead of a subliminal stimulus, an excitation above threshold (if more than one, but less than $a$, $\Theta_i$ were involved at this stage, our argument would only take a few more steps). Then each time *any* triangle $\Theta_i$ is presented to the machine, the common response $\Theta$ (which can convey, clearly, only the "general concept" of triangle) is always evoked, i.e. the neurons of the "common" constellation $c$ fire and their channels are facilitated some more; on the average, the couplings of each constellations $c_i$ will be susceptible to increase for only $1/a$ of the total time, while those of $c$ will be exposed to facilitating actions for *all* the time—until permanent memory is achieved. Even if at the beginning, say, $c_1$ had been markedly facilitated, the presentation of more and more patterns which have "something in common" with it (as decided by the machine) will cause the specific response to $c_1$ to fade into oblivion, while the "abstract", or "common" response is evoked as the first thing, after a convenient learning period.

If we consider now only two stimuli $\Theta_1$ and $\Theta_2$ and assume that $\Theta_1$ *causes* already, say, because of the genetic structure of the machine, *a direct response* (e.g. the food-salivation reflex of Pavlov's dogs), while $\Theta_2$ is ineffective (cf. bell-ringing), then the same mechanism can be clearly extended to account first for the formation of the response $\Theta_{2,1}$ (bell before food, $\neq \Theta_{1,2}$), then, for the fact that $\Theta_2$ alone comes to provoke the same effect as $\Theta_1$ or $\Theta_2$. The temporal behavior given by M.E. of type (5) is perfectly suited to describe the available evidence, which might be used for a determination of the numerical values of some of the constants which appear in (5).

This type of conditioning is manifestly different from that described in 6 above, which requires the intervention of the thalamus: we shall call it *conditioning of the second kind*, or "by information" (the bell just tells the dog that food is coming).

10. We discuss, finally, *re-integration*, which we define as the fact that our machine shall, after learning a pattern, respond to the "incomplete" presentation of that pattern as if it were complete (cf. the familiar oversights of the proofreader). That this must be the case can be seen now quite trivially: after a number of facilitations have occurred, a smaller number of input stimuli will be required to cause the firing of neurons

than were necessary when that pattern was presented for the first time. This also accounts for the children's alterations of new words, which are reduced to combinations of already familiar words; in part, for the fact that a cue suffices to evoke a long string of memories, e.g. verses, etc.

## 5. Concluding Remarks

### A. TIME EVOLUTION OF PATTERN-ANALYSIS

1. The preceding discussion has been, at various places, restricted to situations in which a stimulation, or presentation of a pattern, at the input occurs at given instant, after which it ceases while the machine starts its analysis of it; that is, we have chosen to consider only the "free" modes of motion of the machine, of which the stimulation sets the initial conditions, rather than the evidently prevalent situations in which the machine will perform "forced" motions under the *continued* influence of stimulations which persist and may vary with time.

This simplification is obviously convenient for the purposes of a purely qualitative discussion, as it permits separate examination of the various features of the operation of the machine; nor can forced motions be adequately described without a quantitative analysis of the solutions of the N.E. and M.E. We wish to point out here, however, that this simplifying assumption is, in all likelihood, a much better approximation of reality than it may seem at first.

We can consider a continued stimulation as the presentation to the machine of a *time-series of patterns*; its analysis by the machine is therefore a *serial* operation. Our observation is, that a machine such as the one envisaged by us, with a very large number of elements, can actually transform that serial operation into a *parallel* operation.

Let the patterns presented at the input (or anywhere by the machine to itself: cf. Section **4**, C, 8) at time $0, \tau, 2\tau, 3\tau \ldots$ be $\Theta_0, \Theta_1, \Theta_2 \ldots$. We recall (Section **3**, B, 2) that each neuron, as described by the N.E., has a *refractory* period which is $> \tau$, and may be $\gg \tau$; we take it here to be $R\tau$, according to the schematization expressed by (4). $\Theta_0$ causes a set $S_0$ of neurons to fire at time 0; all neurons of $S_0$ then stay dead for $R\tau$ sec, while other neurons are excited by them at rather distant places (3rd criterion of stability, Section **4**, C, 4) and then remain dead in turn while stimulating other distant neurons, etc. When $\Theta_1$ is presented at time $\tau$, all neurons of $S_0$ *cannot respond*, and another set $S_1$ *disjoint* from $S_0$ fire and excite likewise distant neurons, etc.; and so on for $\Theta_2$ at time $2\tau$, etc.

All the patterns presented from $t_0$ until $t = R\tau$ are therefore registered *in parallel* by the machine, which can thus, for example, abstract a con-

cept meaning "motion" with the same mechanism by which it abstracts one meaning "triangularity".

The extent to which this happens is determined, in this simple example, by the value of $R$. If the reverberations which are significant for pattern-analysis have periods $< R\tau$, then the simplifying assumption made in Section 4 is quite good. In any case, the conversion of serial into parallel operation, and *vice versa*, will emerge as an obvious and important feature in any quantitative discussion of the N.E.

2. It is convenient at this stage to return briefly to the matters mentioned in Section 4, C, 2, so as to summarize the essentials of the operation of the machine into a few mathematical concepts.

*Our machine learns, by the process of abstraction* (by virtue of the M.E.), *to perform pattern-analysis* (by virtue of the N.E.). This sentence contains all that is most relevant in our theory; mathematically, it can be expressed as follows.

At a fixed time $t$, the N.E. have frozen coefficients: A.L.H. A pattern $\Theta$ presented at or into the machine at $t$ evokes, in general, a very large number of disjoint modes, or reverberations: this is the "Fourier" analysis performed by the machine, each mode corresponding to a point $\mathcal{Q}$ or an axis in an appropriate functional space $\mathcal{F}$ on $\mathcal{N}$ (the quotes call to mind the profound difference from linear Fourier analysis).

We may also say that each point $\mathcal{Q}$ of $\mathcal{F}$ corresponds to one of the *basic concept*, or *words*, which the machine has learnt until $t$: the N.E. contain, implicitly, all the *knowledge* or *vocabulary* of the machine, in terms of which each new pattern is translated by the machine. (This vocabulary we expect to have a surprisingly different structure from those of Western languages, in which a word is a "point", and a sentence or definition is a "surface constructed point-by-point"; more akin to Chinese or Japanese, in which a word is a "plane" and a sentence or definition a "surface constructed as envelope of planes".)

The second difference from Fourier analysis (the first being non-linearity) is that now the set of fundamental modes, or axes in $\mathcal{F}$, is not *constant*, but *changes in time* because of the M.E. Thus, at a given $t$ a pattern is represented by a point in a given frame $S$ in $\mathcal{F}$ (N.E.); this frame, however, is not fixed, but changes slowly (A.L.H.) in time (M.E.): $S = S(t)$.

This scheme, if correct, has profound implications: for instance, the efficiency of the machine will depend tremendously on the method followed in its education, because to the same external stimulus machines which have been educated differently may offer very quick and simple, or very slow and involved responses.

3. An interesting consequence of our theory is the fact that a machine

thus constructed necessitates periods of rest, or "sleep". Indeed, even giving it a tremendously large number of elements, the long duration of reverberations (and induced motions in general), which goes far beyond that of the actual stimulations, will cause, if stimuli are unceasingly offered to it, an ever increasing cumulation of activity; the possibility of co-existence of disjoint reverberations will become less and less, interference will cause "confusion of ideas" and inability to give correct answers even to familiar questions.

A period of "sleep", i.e. cessation of activity through the suave interaction of the reticular system, will permit the gradual extinction first of the less facilitated, then of all other reverberations, and the fading of semi-permanent memories of relatively short duration as well. A more quick and drastic treatment, such as "electroshock", followed by total quiescence for only the period of time which is required for such semi-permanent memories to disappear, will produce the same effect. This reminds one of some Yogi techniques which are said to achieve, in a relatively short time, the same state of rest which follows a full night's sleep.

4. As a final remark, we wish to point out that it might be interesting, in the light of the present considerations, to attempt an analysis of E.E.G.'s in terms of subharmonics instead of the customary one in terms of harmonics of a suitably chosen frequency. The *amplitude* of the E.E.G. recordings should depend strongly on the mechanism described in 5, A, 1.

### B. SELF-ORGANIZATION INTO RELIABLE OPERATION

1. In the study of any system containing a very large number of interacting elements built with realistic tolerances—let this be the machine which is being discussed here, or a bee-hive, or the whole socio-political and administrative framework of a nation—the central question is certainly whether this system obeys instantaneous and evolutionary laws which guarantee its spontaneous *convergence in time towards more efficient operation*, or whether the system may rather show *erratic* or *divergent* performance. This question was formulated with full clarity by N. Wiener, who, besides emphasizing the vital importance of it, gave also powerful mathematical tools for its investigation; in his treatment the non-linearity of the interaction laws was rightly stressed as the key to the whole problem.

We can do nothing better than refer the reader to his work for a deeper elaboration of these ideas; it is important, however, here, to show that there is very satisfactory qualitative evidence that a machine obeying *suitable* N.E. and M.E. will satisfy Wiener's principle of self-organization. We shall keep this discussion at a qualitative level by resorting to physical more than to mathematical arguments, and by examining later in detail,

instead of the actual machine proposed here, a simplified functional model of it.

That non-linearity of some sort is necessary for self-organization is physically obvious. A linear system (with *frozen* coefficients, or else non-linearity intervenes: cf. A, 2 above) cannot perform a spontaneous transition from one of its states to another (we use the language of quantum mechanics only because of its greater appeal to intuition; of course, the same is true classically); such transitions—without which the system could not choose spontaneously, from our point of view, states with a better (or worse) organization—are only possible if there are *perturbations* to cause them (e.g. the interaction with the electromagnetic field causes the quantum jumps in atoms); furthermore, *the system itself* must originate these perturbations (the electrons of the atom are the source of the electromagnetic field in spontaneous emission), or else its changes of state would be "induced" rather than "spontaneous". The equations of the system must therefore be non-linear in an essential way; such, for instance, as we have in electrodynamics when the electromagnetic field is expressed in terms of electron sources.

Only a non-linear system of this sort can change its state spontaneously; it will not be generally true, however, that its changes are necessarily "for the better". Again, simple examples suffice to prove this statement; leaving aside those, plentiful indeed, which come from societies or civilizations that go bankrupt, we observe that the behavior of a non-linear physical system the energy of which is not restricted by a finite lower bound is certainly catastrophic. We expect therefore that such a system, in order to satisfy Wiener's principle, shall obey additional *necessary* "convergence" conditions—of the type, for instance, that in quantum mechanics secures the existence of a ground state, which we may identify with the "best state for efficient operation". It is our belief that these conditions will amount to satisfying the third criterion of stability (Section 4, C, 4) and to choosing N.E. with the qualitative behavior exhibited by (5).

The study of the restrictions that this criterion would impose on N.E. (2) and M.E. (5) appears to be a rather straightforward problem, which we add to the list of those which we formulated and set aside in Section 4 for future investigation. A dynamical interpretation of (2) and (5) would be indeed quite natural, as the M.E. (5) just express a non-linear coupling of the N.E. (2) with themselves; we deem it more meritorious, at this stage, to resist the temptation of adapting the available quantum-field-theoretical knowledge to these problems, than to yield to it.

Clearly, the A.L.H. imposes the distinction between two types of non-linearity: that which is expressed by the N.E. (2) and that which is

expressed by the M.E. (5). The first is useful, but only the second is necessary for the type of self-organization we are discussing; this statement will appear obvious after the discussion in 2 below, which is dedicated, as was announced, to a simplified version of our problem.

2. We wish to discuss here a *model* of our machine (which we may also regard as another, though less sophisticated, model of the brain; as such it was briefly discussed at an early stage of our work) which consists in a system of $N$ linearly coupled harmonic oscillators, the constants of which obey M.E. of type (5) ($r = o$), under the A.L.H. This was mentioned in Section **4**, B, 3, and is a model of our machine in the sense that, as was said there, "reverberations" correspond in it to normal modes, "subharmonics" to ordinary harmonics, etc. The wealth of solutions of N.E. of type (2) is now lost, because of the linearity of the N.E. of this model; for it the central problem reduces to the determination of the solutions of a secular equation of degree $N$:

$$f(a_{hk}; \lambda) = 0 \tag{6}$$

This digression is useful both because it readily provides a qualitative insight into the behavior of a learning machine with respect to self-organization, and because it shows that many different mechanisms may be built, with various degrees of convenience, to produce "thought": the only essential thing is that they obey *some* N.E., *some* M.E. (*suitable*, but not necessarily of type (2) and (5) ), and closely enough the A.L.H.

We wish to compute the average change $< \delta\bar{\lambda} >$ of a solution $\bar{\lambda}$ of (6) when the coefficients $a_{hk}$ of the equations of the system undergo infinitesimal random variations (from a "macroscopic" point of view, small variations due to learning will appear as "random"):

$$a_{hk} \to a_{hk}(1 + \rho_{hk}) = a_{hk} + \delta a_{hk} \tag{7}$$

such that $< \rho_{hk} > = \rho$.

From:

$$f(a_{hk} + \delta a_{hk}; \bar{\lambda} + \delta\bar{\lambda})$$

$$= \sum_{hk} a_{hk}\rho_{hk} \frac{\partial f}{\partial a_{hk}} + \left(\frac{\partial f}{\partial \lambda}\right)_{\lambda = \bar{\lambda}} \delta\bar{\lambda} = 0 \tag{8}$$

we find, on taking averages (since (6) is homogeneous of degree zero in the term $\lambda^N$ and of degree one in all other terms, with respect to the variables $a_{hk}$):

$$< \delta\bar{\lambda} > = \frac{(\bar{\lambda})^N}{f'(\bar{\lambda})} \rho \tag{9}$$

which tells us several interesting things. It requires, first of all, that $f'(\bar{\lambda}) \neq 0$, or else $\bar{\lambda}$ would be a degenerate eigenvalue, and even with

$\rho = 0$ the degeneracy might be removed just the same by (7). We suppose therefore $f'(\bar{\lambda}) \neq 0$; to gain some insight into the behavior of $< \delta\bar{\lambda} >$ we just suppose here that the roots of (6) are all simple and equally spaced, so that

$$\bar{\lambda}_k = k\lambda_0, \quad (k = 1, 2, 3, \ldots N) \tag{10}$$

then (9) gives:

$$\left| \frac{< \delta\bar{\lambda}_k >}{\bar{\lambda}_k} \right| = \frac{1}{(N-1)!} \binom{N-1}{k-1} k^{N-1} \rho \tag{11}$$

whence, for $k = 1$ and $k = N$ ($N \gg 0$):

$$\left| \frac{< \delta\bar{\lambda}_1 >}{\bar{\lambda}_1} \right| = \frac{1}{(N-1)!} \rho \tag{12}$$

$$\left| \frac{< \delta\bar{\lambda}_N >}{\bar{\lambda}_N} \right| \sim e^N \rho \tag{13}$$

For our qualitative purposes, (12) and (13) suffice amply to show that random variations, due to learning as expressed by M.E. of type (5), may alter by a vanishing amount the smaller, by increasingly relevant amounts the larger eigenfrequencies of the normal modes of this model.

If a frequency which is kept long enough unchanged becomes permanent, this model will evolve therefore with learning so as to preserve the modes with small frequency; the modes with higher frequency, as well as those that correspond to degenerate solutions, will change at random without staying long enough at any given value to become permanent. After a "long" time, the system will have shifted spontaneously, because of its learning ability, toward a "ground state", in which there is no degeneracy, but the allowed frequencies stay as close to one another as the maximum and minimum values conceded by the M.E. to the constants will permit; this is the "senile age" of the system, in which no learning is possible because all available memories are already engrammed permanently (cf. (7): all $\delta a_{hk} \equiv 0$). The "infancy" of the system is characterized instead by $\rho > 0$, because for $t > 0$ (when all the $a_{hk}(0)$ have minimum absolute values) they can only increase monotonically: learning is somewhat slower, degeneracies are removed faster. The "adult" age corresponds to the period in which it is mostly $\delta a_{hk} \neq 0$, $\rho = 0$.

This cursory and incomplete glance at the properties of this model is intended only to show in which sense we should expect Wiener's principle to be verified by it: the "reliable" information is that carried by the small eigenfrequencies; when the erratic behavior of the higher frequencies pushes one of them down enough, it may be permanently "engrammed", increasing thus the reliability of the system.

3. The main differences between the machine we study and the model of it discussed in 2 above lie, as regards the validity of Wiener's principle, in the non-linearity of the N.E. (2) and in the 2nd criterion of stability (Section **4**, C, 4). The situation is made worse to some extent than in (2) by this non-linearity; in any case, though, by assuming the number of elements to be large enough, one ought to obtain a preferential decrease of the effect of random learning variations on the faster reverberations (to higher frequencies of the model described in 2 slower reverberations of the machine now correspond), so as to reproduce, although perhaps less dramatically, a situation like that represented in 2 by (12) and (13).

The existence of a thalamus and of a reticular system (2nd criterion of stability) makes instead a tremendous improvement upon the model discussed in 2, which could still satisfy Wiener's principle even though it is not endowed with these homeostatic devices. Conditioning of the first kind and arousal of attention by the thalamus, prevention of too diffuse (and therefore not very meaningful) reverberations by the reticular system, are only examples, clearly, of what controlling devices of such effectiveness may do in the way of forcing the machine into learning important information in a reliable manner. These devices are especially important if the memory is prevalently of *negative*, rather than of *positive* type as is assumed here (Section **3**, C, 1), because then the machine would have, in its infancy, a tendency towards epilepsy.

Of importance in a discussion of this principle are also, clearly, all the additional special-purpose devices that Nature uses. The machine also is better with these devices, as intermediate links between external inputs and outputs, and inputs and outputs to the machine proper or cortex; their consideration, however, belongs to engineering more than to physics, and would not be relevant at this place. In conclusion, we should like to stress once more our firm conviction that the speediest way to progress in all the problems connected with the actual construction of machines of this sort is humble resort to Nature's own doings, through neuroanatomy and physiology.

### C. FURTHER OUTLOOKS

As a final comment, we think it appropriate to remark that the general formalism of N.E., M.E. and A.L.H. which is expounded here seems to us to admit of a far wider range of applicability than that to which it has been restricted in this work. The N.E. in fact, for instance, serve only to express, in a more or less schematic manner, the fact that a decision is taken, after a weighted evaluation of the information which lasts a finite time, by a member of a set; and that such a decision is bound to affect other decisions, etc. We may change their name into that of *decision equations*, call the

M.E. *evolution equations*, and take all our considerations over to the study of social or economical or other collective phenomena.

We have pursued this line of thought in several directions for personal amusement, and have soon found, to our surprise, that the qualitative analysis given here for thought-processes applies as well, *mutatis mutandis* (that is, names), to a great many other instances. We believe that, as soon as it becomes possible to agree on a concrete choice of schemes and numbers, quite reasonable predictions may be made in this way about, say, the operation of a stock-exchange, the variation in time of a parameter in feminine fashion, the type of national government that would best obey Wiener's principle, and so on. This we say with at least the same degree of assurance that we have found in the economists who apply the Schrödinger equation to the study of their problems.

Although we have refrained here from a quantitative analysis of the several mathematical problems formulated in the course of this work, the results we have already obtained in this direction seem to justify some optimism; if these expectations are not illusory, then the present formalism might help us to gain a finer knowledge of some physical phenomena that can now be treated only with statistical methods.

This research would not have been possible without the generous and enthusiastic collaboration of Dr. V. Braitenberg, neuroanatomist, among whose merits were certainly not least patience toward this writer's initial ignorance and presumption, and success in eliminating the latter; of Dr. F. Lauria, a young mathematician who dared to place mathematics not too high above common brains; and of many others, to all of whom it is our duty and pleasure to extend our sincerest thanks.

## REFERENCES

Beurle, R. L. (1956). *Phil. Trans.* **B669,** 55.

Braitenberg, V. & Lauria, F. *Nuovo Cim.* Suppl. (In press).

Braitenberg, V., Caianiello, E. R., Lauria, F. & Onesto, N. (1959). *Nuovo Cim.* **11,** 278.

McCulloch, W. S. & Pitts, W. (1943). *Bull. math. Biophys.* **5.**

Moruzzi, G. & Magoun, H. W. (1949). *Electroenceph. clin. Neurophysiol.* **1,** 455.

Wiener, N. (1958). "Nonlinear Problems in Random Theory". Massachusetts Institute of Technology.

*J. Theoret. Biol.* (1961) **2**, 236–243

# A Theoretical Mode of Action of Aldosterone†

Arthur J. Vander, Walter S. Wilde and Richard L. Malvin

*Department of Physiology, School of Medicine, The University of Michigan, Ann Arbor, Michigan, U.S.A.*

(*Received* 10 *January* 1961)

Aldosterone is the most important of known sodium-regulating hormones, although it is responsible for only a few per cent of total renal sodium reabsorption. In addition, aldosterone can increase renal excretion of hydrogen and potassium ions. That its renal effects are direct was first demonstrated by Barger, Berlin & Tulenko (1958) by unilateral renal artery injection of the hormone into normal and adrenalectomized dogs. Moreover, these investigators were able to demonstrate a complete dissociation of aldosterone's kaluretic and anti-natriuretic effects in the normal dog. Ganong & Mulrow (1958), using a similar technic, have also found a significant separation of these effects in the adrenalectomized dog. Thus the possibility that aldosterone accelerates transport systems exchanging sodium for hydrogen and potassium remains an unresolved question.

Recently (Vander *et al.*, 1958; 1960), with the use of stop flow analysis technic (Malvin, Wilde & Sullivan, 1958), a distal site of action of aldosterone has been demonstrated. It is the purpose of this paper to describe a model for sodium transport and aldosterone activity based upon these findings.

## Site of Action of Aldosterone

By administration of an osmotic diuretic (mannitol) to the dog, very high rates of urine flow can be established. If during this period the ureter is clamped, intratubular pressure rises to equal net filtration pressure, at which point glomerular filtration ceases. During this period of stopped ureteral flow, the concentration of any substance in the intratubular fluid column will be changed along the nephron, depending upon how the individual segments handle this substance. The mannitol retards water reabsorption and provides a menstruum against which electrolyte concentrations may be changed. Upon release of occlusion, these concentra-

† Supported by the American Heart Association, 57-G-157; the Life Insurance Medical Research Fund, G-59-44; the National Institute of Arthritis and Metabolic Diseases, U.S. Public Health Service, A-1740-C3S1 Physiology.

tion patterns are obtained in approximately 0·5 ml urine samples which segment this pattern into an ordered array, best pictured on a graph if concentration of each sample is plotted against accumulative urine volume. Thus the dashed curve in Fig. 1 is a typical normal sodium concentration pattern, demonstrating a fall in sodium concentration as fluid trapped in the distal tubules during occlusion enters the urine collector after release of occlusion. The line designating PAH maximum indicates the collection of fluid trapped primarily in the proximal tubules during occlusion.

It is evident from Fig. 1 that, after adrenalectomy, the distal tubule was not able to reduce sodium concentration to the minimum value achieved during ureteral occlusion in normal dogs. The final concentration achieved was not altered by prolonging the length of occlusion.

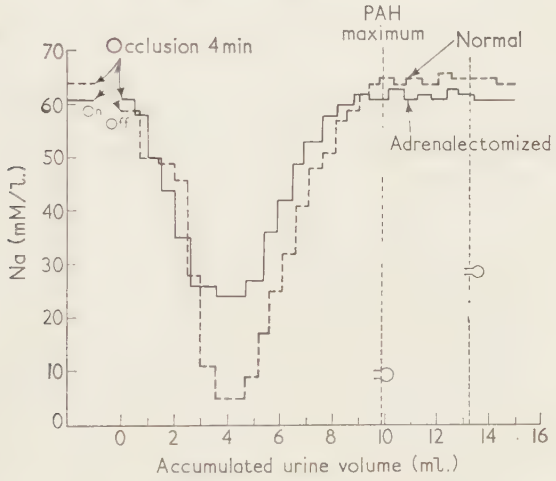



Fig. 1. Comparison of distal tubular sodium reabsorption during ureteral occlusion in a normal and in an adrenalectomized dog (Vander *et al.*, 1958: by courtesy of the Society for Experimental Biology and Medicine).

Also of significance was the finding (Fig. 2) that the minimum distal stop flow sodium concentration was a function of the plasma sodium concentration in the adrenalectomized but not the normal dogs. In the former group, as plasma sodium rose (this was effected by injection of hypertonic NaCl between occlusions), the minimal sodium concentration attained during ureteral occlusion also rose. This indicates that adrenalectomy has reduced the maximal sodium concentration gradient which could be maintained across distal tubular cells. Identical results were obtained by administration of the steroidal antagonist SC-8109 to normal dogs. Finally, all the above abnormalities could be corrected by the administration of aldosterone.

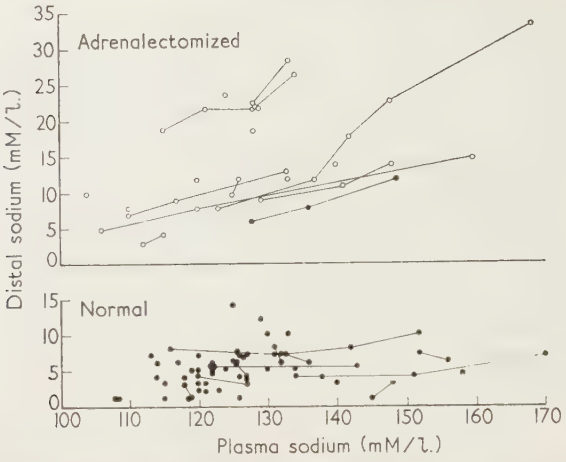



FIG. 2. Relationship between plasma sodium concentration and the minimal distal urinary sodium concentration developed during ureteral occlusion in normal and adrenalectomized dogs. Lines connect points derived from a single dog as plasma sodium was elevated (Vander *et al.*, 1960: by courtesy of the Society for Experimental Biology and Medicine).

## Mode of Action of Aldosterone

Conventional clearance methods indicate that aldosterone increases tubular reabsorption of sodium. This action is manifested in stop flow analysis by the ability of aldosterone to increase the maximal sodium concentration gradient which can be developed between plasma and distal tubular urine. Since the minimal distal sodium concentration attained during ureteral occlusion is independent of the duration of occlusion, this concentration must be an equilibrium value, i.e. the concentration at which sodium movement out of the lumen is equal to sodium movement inward. Aldosterone could act, therefore, in one of two different ways: it could activate carrier systems responsible for sodium transport outward or it might decrease passive back-diffusion of sodium from interstitial fluid into distal tubular lumen.

A rough estimate of such back-diffusion of Na is being made in our laboratory, using the isotope $^{24}Na$. The ureteral occlusion stops filtration so that glomerular substances such as inulin will not enter the concentration pattern except as new filtrate. $^{24}Na$ injected intravenously after the occlusion, by crossing the tubular epithelium transmurally, enters the urinary pattern ahead of inulin. Precise delineation of rates of movement into the stop flow pattern is complicated by the continuing decline of the precursor $^{24}Na$ in the blood plasma during the short two-minute period allowed for the $^{24}Na$ to enter the tubule lumen. The shape of the stop flow

curves for $^{24}Na$ seems unchanged after adrenalectomy or SC-9109 with no suggestion of any increased rate of back-diffusion.

### Effect on Kinetics of Active Transport

As mentioned above, the length of time of occlusion beyond 3 to 4 minutes does not alter the final sodium concentration achieved by the distal tubule during occlusion, indicating that this concentration represents an equilibrium value at which movement of sodium outward is equal to its movement inward. This fact allows us to develop the following equations, which can be used to characterize, subject to the limitations described below, not only the carrier systems responsible for active sodium transport but also the influence aldosterone exerts upon them.

$$D_{in} = K\,Na_p \tag{1}$$

$$D_{out} = K\,Na_i \tag{2}$$

where D = rate of passive diffusion of sodium inward or outward; K = diffusion constant for Na through distal tubular cells in either direction; $Na_p$ = plasma sodium concentration; $Na_i$ = minimum distal tubular sodium concentration attained during ureteral occlusion.

It must be pointed out, however, that the rate of these passive fluxes will be determined not only by the diffusion constant but also by the transtubular electrical potential gradient. Since this electrical gradient will oppose sodium flux in one direction and favor it in the opposite, depending upon the orientation of the charge, it cannot be incorporated into the parameter (K). Giebisch (1958) using the technique of micropuncture, has made four determinations (27, 37, 39, 40 mV, inside negative to outside) of distal transtubular potential in *Necturus*. Interpolation of these data to the dog under conditions of stop flow would be unwarranted. Furthermore, since the origin of the potential is unknown, it is not possible to predict the effects of increasing plasma sodium concentration upon it. Because of the impossibility, at present, of incorporating the electrical potential into the equations below, the authors believe that it would still be of value to attempt to set up a working hypothesis, ignoring this factor, and subject always to question and modification.

If $r_{Na}$ = rate of active reabsorption of sodium, then at equilibrium,

$$K\,Na_p = K\,Na_i + r_{Na}$$

$$K(Na_p - Na_i) = r_{Na} \tag{3}$$

The rate, $r_{Na}$, may be described by an equation analogous to the Michaelis-Menten equation of enzyme kinetics:

$$r_{Na} = \frac{R_{Na}\,Na_i}{Na_i + K_M} \tag{4}$$

Where $R_{Na}$ = maximal or saturation rate of distal tubular sodium reabsorption; $Na_i$ is as above; $K_M$ = that sodium concentration in the distal lumen at which the velocity of active outward sodium transport, in analogy to the Michaelis-Menten constant of enzyme kinetics, is equal to $\frac{1}{2}$ the maximal attainable velocity. Combining equations (3) and (4):

$$K(Na_p - Na_i) = \frac{R_{Na}\, Na_i}{Na_i + K_M} \tag{5}$$

In this equation, K, $R_{Na}$, and $K_M$ are all unknown. However, $Na_i$ and $Na_p$ can be determined for two different steady state concentration gradients, in two separate occlusions on the same dog, in one of which $Na_p$ and $Na_i$ are altered by elevating $Na_p$, resulting in an elevation of $Na_i$. These values can then be substituted separately into equation (5) to yield two equations. Divide one equation by the other:

$$\frac{K(Na_p^a - Na_i^a)}{K(Na_p^b - Na_i^b)} = \frac{R_{Na}\, Na_i^a}{Na_i^a + K_M} \Big/ \frac{R_{Na}\, Na_i^b}{Na_i^b + K_M} \tag{6}$$

in which the superscripts *a* and *b* designate sodium concentrations at two different equilibria in the same animal. Since it is necessary that $Na_i$ vary with $Na_p$, such an equation can be obtained only in experiments on adrenalectomized or SC-8109 treated dogs. In the normal animal, this cannot be done, distal sodium concentrations being essentially unchanged

TABLE 1

*Values of* $K_M$ *in adrenalectomized and SC-8109 treated dogs*

| $K_M$ mmoles/l | |
|---|---|
| SC-8109 | Adrenalectomized |
| 5·3 | 4·5 |
| 4·6 | 2·7 |
| 3·3 | 2·5 |
| | 1·7 |
| | 1·6 |

See text for definition of $K_M$.

at all plasma levels of sodium (very small changes, which very likely did occur, cannot be detected by stop flow analysis). The constants K and $R_{Na}$ will cancel out and the equation may then be solved for $K_M$:

$$\frac{Na_p^a - Na_i^a}{Na_p^b - Na_i^b} = \frac{Na_i^a(Na_i^b + K_M)}{Na_i^b(Na_i^a + K_M)} \tag{7}$$

Table 1 lists values of $K_M$ calculated for adrenalectomized and SC-8109 treated dogs. Evidence to be presented below indicates that $K_M$ is identical

in intact animals. It can be seen that $K_M$ is relatively constant in different animals and is quite low. This indicates that distal sodium reabsorptive systems operate at near maximal velocity over a wide range of intratubular sodium concentrations.

The determination of $K_M$, as described above, depends upon the basic assumption that the Michaelis-Menten equations developed do actually apply under the existing experimental conditions. However, it is not at first clear how these might be tested. Absolute velocities cannot be determined,

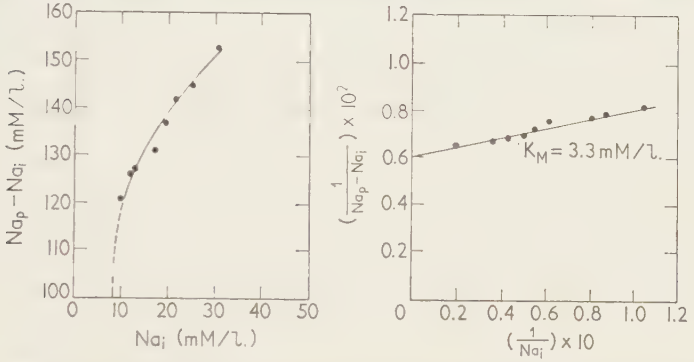



FIG. 3. "Velocity-substrate" plot of data obtained in an SC-8109 treated dog by elevating plasma Na. $Na_p$ = plasma Na; $Na_i$ = distal urinary Na concentration developed during ureteral occlusion. ($Na_p - Na_i$) is the proportional velocity of Na reabsorption. See text for discussion of $K_M$.

nor can they be calculated since a value for the diffusion constant, K, in equation (3) is unknown. Although it would be interesting to evaluate K, the validity of fitting the data to these equations can be determined by another device. This fact can be appreciated by referring back to equation (3). It is important to note again that this equation applies only at equilibrium. It is evident that for any given sodium concentrations in plasma and distal tubular fluid, velocity of sodium reabsorption will be proportional to the difference between the concentrations ($Na_p - Na_i$). If the difference were to be increased by a factor of two, then the relative equilibrium velocity would also be doubled. It becomes evident, therefore, that these concentration differences can be used to represent transport velocities at each of the distal tubular sodium concentrations obtained within the same animal. Different dogs can be compared in this manner only if it is assumed that the diffusion constant, K, is identical in all animals.

Data obtained for an SC-8109 treated animal in which distal tubular sodium concentration achieved during stop flow was changed by increasing plasma sodium concentration are shown as a reciprocal plot in Fig. 3.

These "velocity-substrate" plots indicate that the data really fit a Michaelis-Menten type analysis. The intercept on the vertical axis in Fig. 3 is equal to the reciprocal of the maximal attainable transport velocity, $1/R_{Na}$. If the line is extended to the left, the intercept on the horizontal axis is equal to $-1/K_M$, thus providing a graphical method by which $K_M$ can be determined. It is evident that this plot of reciprocals is merely a repetition of the use of equation (7) and will yield the same values for $K_M$, but only if the data really are linear, i.e. fit the proposed analysis. All values for $K_M$ (see Table 1) derived by equation (7) are identical to those determined independently by this graphical means. The fact that the data are linear may further indicate that the possible error introduced by omitting electrical potentials may not significantly alter results.

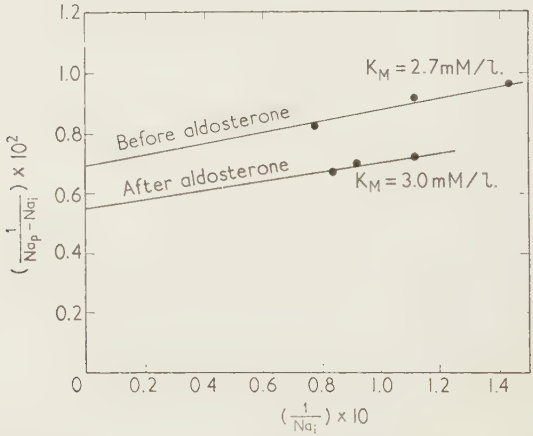



FIG. 4. Effect of aldosterone on carrier systems responsible for distal sodium reabsorption in an adrenalectomized dog. $Na_p$ = plasma Na; $Na_i$ = distal urinary Na concentration developed during ureteral occlusion. $(Na_p - Na_i)$ is the proportional velocity of sodium reabsorption. See text for discussion of $K_M$.

Figure 4 demonstrates the action of aldosterone upon distal sodium reabsorptive systems, as revealed by these graphical methods. Three occlusions were performed in an adrenalectomized dog, increasing plasma sodium after each occlusion. After completion of the third occlusion, 6 $\mu$g aldosterone were administered intravenously over 5 minutes and 1 $\mu$g/hr was incorporated into the normal infusion. Three occlusions were then performed at 15 minute intervals after first waiting 2 hr 10 min. Plasma sodium was again elevated by salt injection between occlusions. Aldosterone dosage was kept at this level so that it permitted a degree of sodium elevation in the distal tubule as plasma sodium was elevated. As already described, this is extremely difficult to do in a normal animal or in an

adrenalectomized animal given large amounts of aldosterone. It should be pointed out that comparison depends upon the assumption that the diffusion constant, K, is not changed by aldosterone, an assumption believed to be valid on the basis of the isotope studies described earlier.

From the data shown in Fig. 4, values for $K_M$, before and after aldosterone, can either be calculated using equations described above, or graphically determined by extending the lines to the left and measuring the intercepts on the horizontal axis $-1/K_M$. It is this constant, a measure of the degree of association between carrier and sodium (in analogy to the affinity of an enzyme for its substrate) that would be altered by a competitive inhibitor or a coupling activator, but it can be seen that $K_M$ is essentially unchanged by aldosterone.

Without changing this intercept on the horizontal axis, aldosterone has decreased the intercept on the vertical axis $1/R_{Na}$, so that the entire "velocity-substrate" line lies under that determined for the untreated adrenalectomized dog. Thus, aldosterone, without altering $K_M$, has increased transport velocity at each substrate concentration as well as having increased the maximal transport velocity attainable at infinite substrate concentration. This is the pattern proposed for a "non-coupling" activator (Friedenwald & Maengwyn-Davies, 1954). Stated more simply, aldosterone does not change the characteristics of the carrier systems, but increases the effective concentration of these carriers available for sodium reabsorption.

Although the data fit the Michaelis-Menten equation, this does not exclude the possibility that some other type of transport system geared to the saturation principle is involved. However, since there is no direct evidence bearing on the exact nature of the Na transport system, one can only speculate as to possible reactions involved. The above analysis merely presents a model for Na transport and describes the influence of aldosterone upon that system.

## REFERENCES


Barger, A. C., Berlin, R. D. & Tulenko, J. F. (1958). *Endocrinology*, **62**, 804.

Friedenwald, J. S. & Maengwyn-Davies, C. D. (1954). "The Mechanism of Enzyme Action", 154-179. (W. D. McElroy and B. Glass, eds.) The Johns Hopkins Press, Baltimore.

Ganong, W. F. & Mulrow, P. J. (1958). *Amer. J. Physiol.* **195**, 337.

Giebisch, G. (1958). *J. cell. comp. Physiol.* **57**, 221.

Malvin, R. L., Wilde, W. S. & Sullivan, L. P. (1958). *Amer. J. Physiol.* **194**, 135.

Vander, A. J., Malvin, R. L., Wilde, W. S., Lapides, J., Sullivan, L. P. & McMurray, V. M. (1958). *Proc. Soc. exp. Biol. Med.* **99**, 323.

Vander, A. J., Wilde, W. S. & Malvin, R. L. (1960). *Proc. Soc. exp. Biol. Med.* **103**, 525.

*J. Theoret. Biol.* (1961) **2**, 244–257

# Replacement of Amino Acids in Proteins

MARTYNAS YČAS

*Department of Microbiology, State University of New York Upstate Medical Center, Syracuse, N.Y., U.S.A.*

(*Received* 13 *January* 1961)

Some consequences of the hypothesis that ribonucleic acid codes amino acid sequences with a coding ratio of one are discussed. The hypothesis predicts certain limitations on amino acid replacements in proteins as a result of mutation. Available evidence is reviewed and found to conform with prediction.

Recently I reviewed evidence indicating that information determining the sequence of amino acid residues in viral protein is coded in viral ribonucleic acid with a coding ratio of one, so that one nucleotide corresponds to and determines one amino acid residue. Part of the evidence for this hypothesis is that amino acids can be assigned to nucleotides in such a manner that the mole fraction of each group of amino acids in viral protein is equal to the mole fraction of the corresponding nucleotide in the viral ribonucleic acid. It is thus possible to predict the composition of viral ribonucleic acid from knowledge of the composition of the corresponding viral protein. Assignments producing this result are as follows: *Adenylic acid:* glutamic acid, glutamine, glycine, leucine, phenylalanine and tryptophan. *Uridylic acid:* aspartic acid, asparagine, isoleucine, histidine and serine. *Guanylic acid:* alanine, arginine, tyrosine and valine. *Cytidylic acid:* cysteine, lysine, methionine, proline, and threonine (Yčas, 1960).

It is the purpose of this paper to develop certain deductions of the hypothesis as they bear on mutations affecting protein structure, and to show that these deductions are supported by empirical evidence.

A coding ratio of one evidently implies that viral ribonucleic acid possesses only part of the information required to specify a protein, since each nucleotide limits, but does not finally determine, the choice of a residue at any given position.† Supplementary information must be contributed by some structure in the host. As illustrated in Fig. 1, the complete information bearing structure is assumed to be composed of two

† For evidence against the supposition that neighboring nucleotides provide additional information see Yčas, 1958.

components: $\alpha$, here viral ribonucleic acid, and $\beta$. The $\beta$ component is as yet chemically unspecified and may itself be a compound structure. On general biological grounds it seems reasonable to postulate that, if this is correct, such a scheme may apply generally: a $\beta$ and a non-viral $\alpha$ component being always present in the cell for purposes of normal protein synthesis.

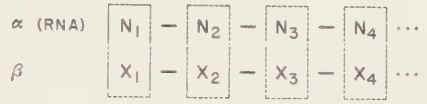



FIG. 1. Illustration of the concept that the material specifying the amino acid sequence in proteins is a double structure. An amino acid at any position is completely specified by a nucleotide in RNA and an element X in the $\beta$ component.

Consider now the result of a mutation in the $\alpha$ or $\beta$ component. (The term "mutation" is here used in the sense of any inherited change, whatever its nature.) To change the $\alpha$ component, at least one nucleotide must be replaced by another of a different kind, and the amino acid previously specified at that position must therefore be replaced by one belonging to a group specified by a different nucleotide. For example, if the original nucleotide was adenylic acid, the new amino acid will be one specified by cytidylic, guanylic or uridylic acid, but not by adenylic acid.

If, on the other hand, a change is produced in the $\beta$ component, the new amino acid will belong to the group determined by the same nucleotide as the one previously occupying that position, since the nucleotide in the $\alpha$ component remains the same. Thus replacements prohibited as a result of "$\alpha$ mutation" are the only ones permitted as a result of "$\beta$ mutation", and vice versa. Since the assignments of amino acids to nucleotides are, by hypothesis, known, these predictions are quite definite and can, in principle, be checked by observation.

In the case of viruses, it is possible to be certain that all mutations are $\alpha$ mutations, since different mutants, i.e. $\alpha$ components, can be grown in the same host and thus produce proteins with information contributed by the same $\beta$ component. All amino acid replacements in viral protein resulting from mutation must therefore involve amino acids assigned to different nucleotides. The evidence available does not contradict this requirement (Table 1), but unfortunately it is as yet too meager to be conclusive.

More evidence is available on replacements involving non-viral proteins. As in these cases it is not possible to separate experimentally the $\alpha$ and $\beta$ components, it might seem hopeless, at first sight, to observe separately the processes of $\alpha$ and $\beta$ mutations. However, there remains the possi-

TABLE I

*Known sequences of homologous proteins. Replacements differentiating "recently divergent" proteins boxed in. The references in some cases are not to the original work, but to convenient compilations. Amino acid abbreviations as in Fig. 2.*

*Insulins*

| | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | Sei Whale (Ishihara *et al.*, 1958) | Ala | Ser | Thr | | | | | | | | | | | |
| | | | | | | | | Sperm Whale | Thr | Ser | Ilu | | | | | | | | | | | |
| | | | | | | | | Horse | Thr | Gly | Ilu | | | | | | | | | | | |
| | | | | | | | | Sheep | Ala | Gly | Val | | | | | | | | | | | |
| | | | | | | | | Pig | Thr | Ser | Ilu | | | | | | | | | | | |
| A Chain† (Sanger, 1960) | H.Gly | Ilu | Val | Glu | Gln | Lys | Cys | Ala | Ser | Val | Cys | Ser | Leu | Tyr | Gln | Leu | Glu | Asn | Tyr | Cys | Asn.OH | |
| B Chain‡ (Sanger, 1960) | H.Phe | Val | Asn | Gln | His | Leu | Cys | Gly | Ser | His | Leu | Val | Glu | Ala | Leu | Tyr | Leu | Val | Cys | Gly | Glu | Arg |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 |

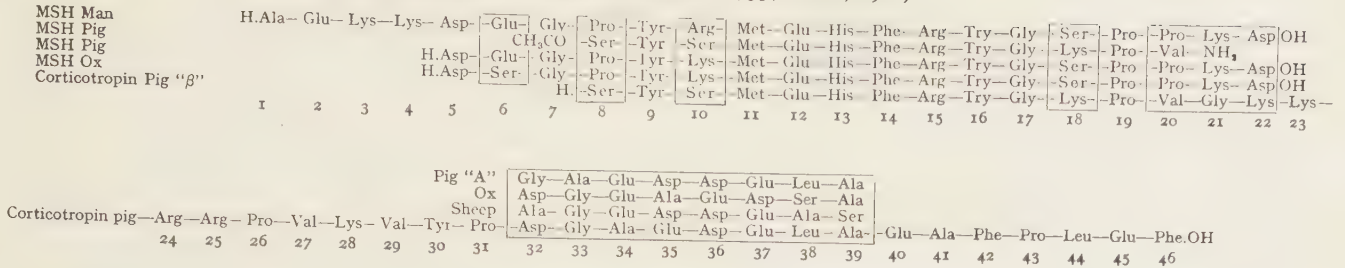
| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | Rabbit | Ser.OH | | |
| | | | | | | | Man | Thr.OH | | |
| Gly | Phe | Phe | Tyr | Thr | Pro | Lys | Ala.OH | | | |
| 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | | | |

*B Chain, Bonito* (Yamamoto *et al.*, 1960)

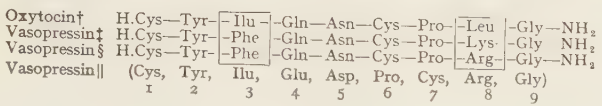
Insulin I H.Leu .....
Insulin II H.Ala ..... } Pro—Lys.OH

† Cattle, Man, Rabbit
‡ Whale, Horse, Sheep, Pig, Cattle

*Melanophore-stimulating hormones and corticotropins* (White & Landman, 1955; Acher, 1960)

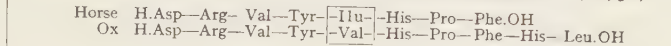



*Vasopressins, Oxytocins* (Du Vigneaud *et al.*, 1953a, b; Acher, 1960)




† Man, Cattle, Pig, Horse, Sheep, Chicken
‡ Pig
§ Man, Cattle, Horse, Sheep
‖ Frog (Acher *et al.*, 1960)

*Hypertensive peptide* (Elliot & Peart, 1956; Skeggs *et al.*, 1956)



*Cytochromes C* (Tuppy, 1959)

Hen ...Val—Gln—Lys—Cys—[-Ser-]-Gln—Cys—His—Thr—Val—Glu...
Salmon ...Val—Gln—Lys—Cys—[-Ala-]-Gln—Cys—His—Thr—Val—Glu...
Cattle, Horse, Pig...Val—Gln—Lys—Cys—[-Ala-]-Gln—Cys—His—Thr—Val—Glu—Lys...

Silkworm ...Val—Gln—Arg—Cys—Ala—Gln—Cys—His—Thr—Val—Glu...
Yeast ...Phe—Lys—Thr—Arg—Cys—Glu—Leu—Cys—His—Thr—Val—Glu...
*Rhodospirrillum Rubrum* ...Cys—Leu—Ala—Cys—His—Thr—Phe—Asp—Glu—Gly—Ala—Asp—Lys...

---

*Serum albumins* (Thompson, 1954; Peters *et al.*, 1958)

Man H.Asp-[-Ala] ....Gly—Val—Ala—[Leu.OH]
Cattle H.Asp-[-Thr] .....[Ala.OH]
Horse H.Asp.....Ala—[Leu.OH]

Chicken H.Asp.....[.Ala.OH]
Turkey H.Asp.....[.Val.OH]

*Growth hormones* (Parcells & Li, 1958; Li *et al.*, 1958)

Man [H.Phe—Ser—Thr.][.....Leu-]-Phe.OH
Humpback Whale [H.Phe.....][.Leu—Ala-]-Phe.OH
Sheep† } { [H.Phe—Thr—Ala.][.Ala—Leu-]-Phe.OH }
Cattle† } { [H.Ala—Phe—Ala.][.Ala—Leu-]-Phe.OH }

† 2 chains

---

*Protamines*

Salmine (Monier & Jutisz, 1954) H.Pro-[-Arg-]-Arg...
Clupein (Waldschmidt-Leitz, 1953) H.Pro-[-Ser-]-Arg...
1 2 3

*Salmo irrideus* (Felix, 1952) ($Gly_2$, $Ser_2$, $Ala_2$, [$Val_4$, $Ilu_1$])†
*Salmo trutta* (Felix, 1952) ($Gly_3$, $Ser_2$, $Ala_2$, [$Val_5$, $Ilu_0$])†

† Composition, not sequence

*Fibrinogens* (Lorand & Middlebrook, 1953)

Man { H.Tyr... / [H.Ala...] }
Cattle { [H.Tyr...] / [H.Glu...] }

*Ribonuclease* (Hirs *et al.*, 1960)

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Sheep | H.Lys | Glu | Ser / Thr | Ala | Ala | Ala | Lys | Phe | Glu | Arg | Ser | Thr | Ser | Ser | Asn | His | Met | Glu | Ala | Ala | Ser | Ser | Ser | Asn | Tyr | Cys | Asn | Gln | |
| Cattle | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | |

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Met | Met | Lys | Ser | Arg | Asn | Leu | Thr | Glu / Lys | Asp | Arg | Cys | Lys | Pro | Val | Asp | Thr | Phe | Val | His | Glu | Ser | Leu | Ala | Asp | Val | Gln | Ala | Val | |
| 29 | 30 | 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 | 41 | 42 | 43 | 44 | 45 | 46 | 47 | 48 | 49 | 50 | 51 | 52 | 53 | 54 | 55 | 56 | 57 | |
| Cys | Ser | Glu | Lys | Asn | Val | Ala | Cys | Lys | Asn | Gly | Thr | Asn | Gln | Cys | Tyr | Gln | Ser | Tyr | Ser | Thr | Met | Ser | Ilu | Thr | Asp | Cys | Arg | Glu | Ser |
| 58 | 59 | 60 | 61 | 62 | 63 | 64 | 65 | 66 | 67 | 68 | 69 | 70 | 71 | 72 | 73 | 74 | 75 | 76 | 77 | 78 | 79 | 80 | 81 | 82 | 83 | 84 | 85 | 86 | 87 |
| Thr | Gly | Ser | Lys | Tyr | Pro | Asn | Ala | Cys | Tyr | Lys | Thr | Thr | Asn | Ala | Gln | Lys | His | Ilu | Ilu | Val | Ala | Cys | Glu | Gly | Asn | Pro | Tyr | Val | Pro |
| 88 | 89 | 90 | 91 | 92 | 93 | 94 | 95 | 96 | 97 | 98 | 99 | 100 | 101 | 102 | 103 | 104 | 105 | 106 | 107 | 108 | 109 | 110 | 111 | 112 | 113 | 114 | 115 | 116 | 117 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Val | His | Phe | Asp | Ala | Ser | Val.OH |
| 118 | 119 | 120 | 121 | 122 | 123 | 124 |

*Tobacco mosaic virus protein*

TMV (Anderer *et al.*, 1960) Acetyl—
TMV (Tsugita *et al.*, 1960) Acetyl—

—Ser—Tyr—Ser—Ilu—Thr—*Pro*—*Thr*—Ser—Gln—Phe—Val—Phe—Leu—Ser—Ser—Ala—Try—Ala—Asp—Pro—Ilu—*Glu*—Leu—Ilu—*o*—Asx (Cys, Thr, Asx)-*Leu*-Ala
—Ser—Tyr—Ser—Ilu—Thr—*Thr*—*Pro*—Ser—Gln—Phe—Val—Phe—Leu—Ser—Ser—Ala—Try—Ala—Asp—Pro—Ilu—*Gln*—Leu—Ilu—Leu—Asp Cys—Thr—Asn—*o*—Ala

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30

—Leu—Gly—Asn—Gln—Phe—Glx—Thr—*Glu*—Gln—Ala—Arg—Thr—Val—Glx—Val—Arg—Gln—Phe—Ser—Gln—Val—Try—Lys—Pro—*Pro*—*Ser*—Gln—Val—Thr—Val
—Leu—Gly—Asn—Gln—Phe—Gln—Thr—*Gln*—Gln—Ala—Arg—Thr—Val—Gln—Val—Arg—Gln—Phe—Ser—Gln—Val—Try—Lys—Pro—*Ser*—*Pro*—Gln—Val—Thr—Val

31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60
31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60

—Arg—Phe—Pro—Asx—Ser—Asp—Phe—Lys—Val—Tyr—Arg—Tyr—Asx—Ala—Val—*o*—Asp—Pro—Leu—Val—Thr—Ala—Leu—Leu—*Leu*—Gly—Ala—Phe—Asx—
—Arg—Phe—Pro—Asp—Ser—Asp—Phe—Lys—Val—Tyr—Arg—Tyr—Asn—Ala—Val—*Leu*—Asp—Pro—Leu—Val—Thr—Ala—Leu—Leu—*o*—Gly—Ala—Phe—Asp—

61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88
61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88

—Thr—Arg—Asn—Arg—Ilu—*o*—Glx—Val—Glx—Asx (Glx, Ala, Asx, Pro, Thr) Thr—Ala—Glx—Thr—Leu—Asx—Ala—Thr—Arg—Arg—Val—Asx—Asx—Ala—
—Thr—Arg—Asn—Arg—Ilu—*Ilu*—Gln—Val—Glu—Asn—Gln—Ala—Asn—Pro—Thr—Thr—Ala—Glu—Thr—Leu—Asp—Ala—Thr—Arg—Arg—Val—Asp—Asp—Ala—

89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116
89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117

—Thr—Val—Ala—Ilu—Arg—Ser—Ala—Asx—Ilu—Asn—Leu—Ilu—Val—Glu—Leu—*Ilu*—Arg—Gly—Thr—Gly—Ser—Tyr—Asn—Arg—Ser—Ser—Phe—Glx—Ser—Ser—
—Thr—Val—Ala—Ilu—Arg—Ser—Ala—Asp—Ilu—Asn—Leu—Ilu—Val—Glu—Leu—*Leu*—Arg—Gly—Thr—Gly—Ser—Tyr—Asn—Arg—Ser—Ser—Phe—Glu—Ser—Ser—

117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146
118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147

Mutant (Tsugita & Fraenkel-Conrat, 1960) Leu
Strain HR
(Niu & Fraenkel-Conrat, 1955) (Thr, Ala)
—Ser—Gly—Leu—Val—Try—Thr—Ser—Gly—Pro—Ala—Thr.OH
—Ser—Gly—Leu—Val—Try—Thr—Ser—Gly—Pro—Ala—Thr.OH

147 148 149 150 151 152 153 154 155 156 157
148 149 150 151 152 153 154 155 156 157 158

Sequence of Tobacco Mosaic Virus protein as determined in two different laboratories. Blank spaces (*o*) inserted where necessary to preserve homology of sequence. Differences in reported sequences in *italics*.

## *Hemoglobins*

| | 1 | | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | | | 22 |
| **Man,** α chain (Braunitzer *et al.*, 1960b) | H.Val— | o | -Leu- | Ser- | -Pro— | Ala- | Asp | Lys | Thr- | Asp | Val | HbJ (Braunitzer *et al.*, 1960c) Lys | Ala | Ala | Try | Gly | Asn Lys | Val | Gly | Ala- | His | Ala | Glv | Glu |
| **Man,** β chain (Braunitzer *et al.*, 1960b) | H.Val | His- | -Leu- | Thr- | Pro- | -Glu- | Glu | Lys | Ser | Ala | Val | -Thr- | Ala | Leu | Try | Gly | Lys | Val | Asp | Val— | Asp | o | o | Glu |
| **HbS (Hunt & Ingram, 1959)** | H.Val— | His- | -Leu- | Thr- | Pro- | -Val- | Glu- | -Lys | | | | | | | | | | | | | | | | |
| **HbC (Hunt & Ingram, 1960)** | H.Val— | His- | -Leu- | Thr- | Pro- | -Lys- | Glu- | -Lys | | | | | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | | | |
| **HbG (Hill & Schwartz, 1959)** | H.Val— | His- | -Leu- | Thr- | Pro- | -Glu- | Gly- | -Lys | | | | | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | | | |
| **Horse** (Porter & Sanger, 1949)β | H.Val— | Gln- | -Leu.... | | | | | | | | | | | | | | | | | | | | | |
| **Cattle** (Porter & Sanger, 1949) | H.Met.... | | | | | | | | | | | | —Tyr— | Gly— | (Ala- | -Glu) | —Ala— | Ser- | -Glu— | Arg— | Met— | Phe... | | |
| | (His- | Val— | Leu- | Leu- | Thr— | Pro— | Glu— | Glu- | Lys)† | | | | -Val | Gly— | Gly | -Glu | —Ala | Leu | Gly | Arg | Leu.... | | | |
| | | | | | | | | | | | | HbE (Ingram, 1959; Braunitzer *et al.*, 1960a) | | | | Lys | | | | | | | | | |

To preserve homology between the α and β chains of human hemoglobin, —o— has been inserted where deletions are presumed to occur.

†Originally published, later corrected sequence in the β chain of human hemoglobin A (Ingram, 1959; Hunt & Ingram, 1959)

---

## *Myoglobins* (Ingram 1955; Kendrew *et al.* 1954)

| | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Horse | H.Gly- | -Leu.... | | | | | | | | | | | | | | | | | | | | | | |
| Finback Whale | H.Val. | .... | | | | | | | | | | | | | | | | | | | | | | |
| Sperm Whale | H.Val. | .... | | | | | | | | | | | | | | | | | | | | | | |
| Seal (*Phoca vitullina*) | H.Gly. | .... | | | | | | | | | | | | | | | | | | | | | | |

---

bility that the $\alpha$ and $\beta$ mutation rates are significantly different, and thus could be distinguished on this basis.

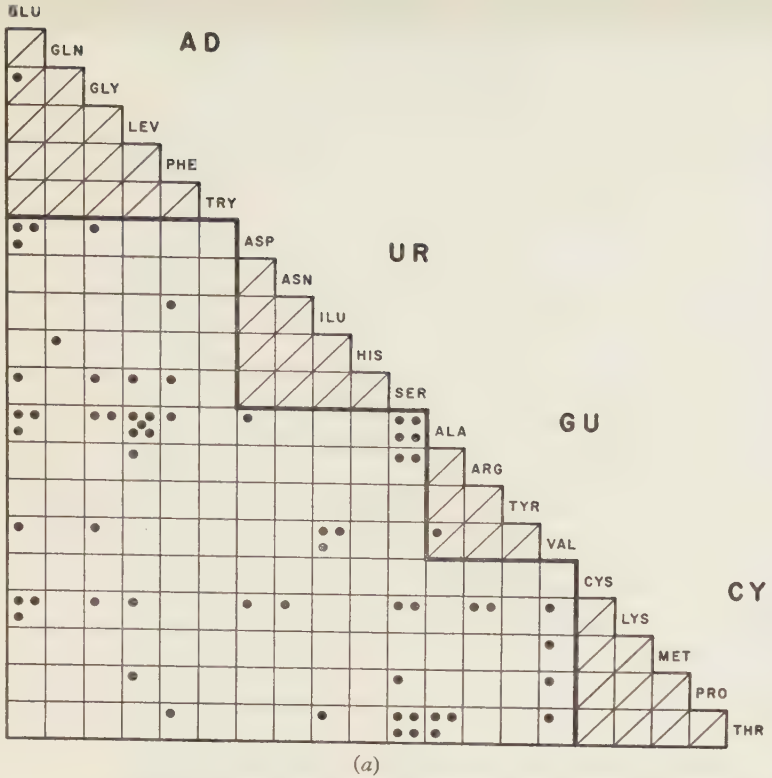
Suppose, as an example, that $\alpha$ mutation is the more frequent. Then replacements arising during some short period of time and differentiating originally identical proteins would be mainly of a type prohibited to $\beta$ mutation. However, if the period of observation were lengthened, $\beta$ mutations would begin to accumulate, so that eventually observed differences between two originally identical proteins would involve replacements which were randomly distributed. Providing only that the rate of $\beta$ mutation were not zero, randomization of differences would always result. The same result would be eventually attained if $\beta$ mutation were more frequent, with, of course, an initial preponderance of differences prohibited by $\alpha$ mutation.

Thus, observation of proteins originally identical and differentiating for varying lengths of time should make it possible to determine if the distinction between $\alpha$ and $\beta$ mutations is real, providing that the rates of these two types of mutations differ.

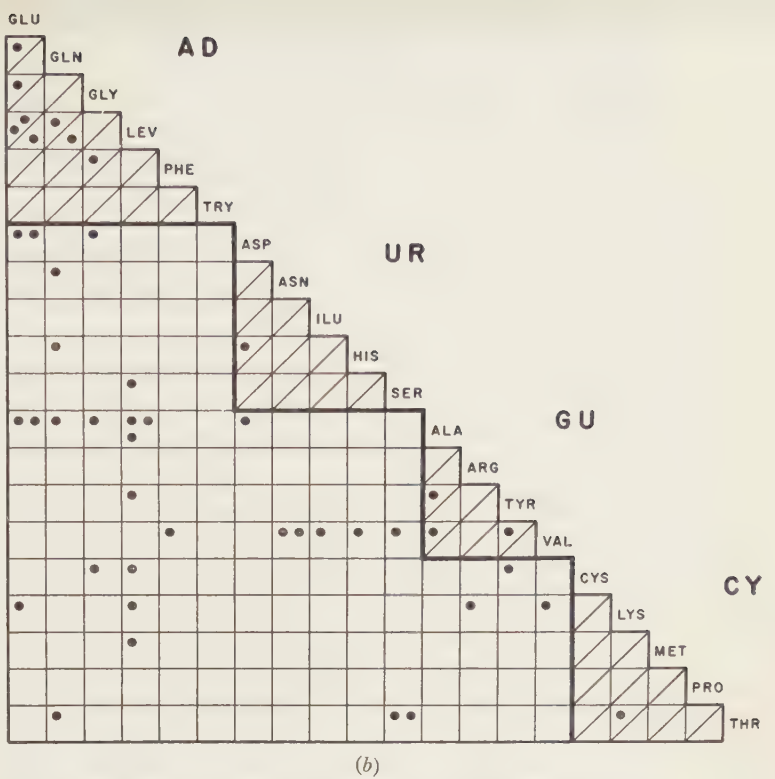
Because of technical difficulties, replacements arising as a result of observed mutations are not known. However, a considerable body of evidence on sequences of amino acids in homologous proteins from different species exists (Table 1). Assuming the truth of the theory of evolution, hereditary differences must be assumed to arise as a result of mutations. The taxonomic distance between species can be used to measure, roughly, the length of time the proteins of two species have been diverging from a common ancestor, so that it is also possible to classify homologous proteins as "recently" and "anciently" divergent. As the rate of change of proteins appears to be very slow "recent" and "ancient" is to be measured in units of geological, rather than ordinary time.

Before comparing protein sequences, it is necessary to clarify the meaning of the term "homologous". In the usual sense it refers to proteins of the same general structure and function in different species that have identical sequences except at a few positions, e.g. insulin from man and pig. Clearcut cases of this type will be referred to as class A homology. However, some proteins, while obviously homologous in the sense of having similar sequences, do not exactly fit this definition. For example, oxytocin and vasopressin have seven of nine amino acids identical, but occur in one organism and have different functions. The same is true of the corticotropins, two forms of which are found in the pig, both similar and having an amino acid sequence closely resembling the melanophore-stimulating hormones. These instances will be regarded, rather arbitrarily, as class A homologous for present purposes.

A homology can also exist between different chains of a single protein




(*a*)




(*b*)

FIG. 2. Distribution of replacements differentiating "recently" (*a*) and "anciently" (*b*) divergent proteins, compiled from Table 1. Hatched areas are replacements prohibited as a result of *a* mutation. AD (adenylic acid), UR (uridylic acid), GU (guanylic acid) and CY (cytidylic acid) are nucleotides specifying the amino acids shown to the left. Amino acid abbreviations as in my 1958 paper except Gln—Glutamine, Asn—Asparagine, Ilu—Isoleucine.

(class B). In a previous publication in collaboration with Gamow and Rich (Gamow, Rich & Yčas, 1956) I have drawn attention to the similarity between the two chains of insulin, as well as to the tendency for the end groups of different chains of single proteins to be identical (Yčas, 1958). As an explanation it was suggested that the several chains were originally one protein specified by a single genetic locus. By reduplication of genetic material the chains began to be controlled by different loci, which gradually diverged by genetic drift. At that time the evidence for this suggestion was slight, but the recent work of Braunitzer, Liebold, Müller & Rudloff (1960b) has, at least in part, vindicated this position. They have shown that the $\alpha$ and $\beta$ chains of human hemoglobin are so similar that a common origin is clearly indicated. As the hemoglobin of the lamprey, presumably representative of the cyclostome ancestor of higher vertebrates, is single chained (Allison *et al.*, 1960), the two chains of hemoglobin must have begun to diverge since about the Devonian or Silurian period, and are an example of homologous proteins long divergent. Since fish insulins are also two-chain proteins (Yamamoto, Kotaki, Okuyama & Satake, 1960), the chains must have diverged at a comparable, probably even earlier period, as might indeed be concluded from the relatively slight similarity remaining between them. Of course the mere presence of two chains must not, by itself, be taken as evidence that the divergence of the chains is ancient. Thus the growth hormone of man is single-chained, while that of cattle is, apparently, double (Table 1). Unless the single chain condition is the result of loss, this double-chain structure must be a relatively recent development.†

On the basis of these conclusions I classify known protein sequences into two groups: recently divergent, mostly mammalian so the divergencies are not older, and preponderantly much younger, than the mesozoic; and anciently divergent, which includes the two chains of insulin and hemoglobin and homologous sequences of cytochrome C from *Rhodospirillum rubrum*, yeast and cattle. These cannot have diverged later than the Devonian, and in the case of the cytochrome certainly vastly earlier.‡

† An examination of homologous sequences in Table 1 shows that, besides the well-known phenomenon of replacement of one amino acid by another, a deletion of residues can also occur, as first pointed out by Šorm and his collaborators (1957). This is clearly demonstrated by the sequences of the two chains of hemoglobin. The known sequences of melanophore-expanding hormones appear to be fragments of an originally longer sequence, which has suffered breakage at various distances from both ends. A C-terminal deletion also seems to have occurred in fish insulin.

‡ Although in principle data on amino acid sequences are reproducible at will, in practice the great labor involved in obtaining them makes an independent verification difficult, so that most of such data must be accepted on the basis of the work of a single laboratory. Some parts of the sequence of ribonuclease, for example, depend indeed on a single experimental determination (Hirs, Moore & Stein, 1960). It is therefore legitimate to ask to what extent such data may be contaminated by experimental error. No general

It is now possible to consider whether there is, in fact, any difference between the distribution of replacements in recently and anciently divergent proteins. Replacements are considered to be all permutations of different amino acids which occupy one position in homologous proteins, listed in Table 1. Thus position 9 in the A chain of insulin is evidence for one replacement, serine-glycine; position 30 in the B chain for three, serine-threonine, serine-alanine and threonine-alanine. If one plots replacements in recently and anciently divergent homologous proteins on a grid, letting the hatched areas represent replacements prohibited as a result of an $\alpha$ mutation, one finds the distributions shown in Fig. 2. If replacements were random, they would be distributed between the two areas approximately in the ratio of 149 : 41.

In fact, replacements differentiating the recently divergent proteins are distributed quite differently. Of 62 such replacements, only two (glutamic-glycine, hemoglobins A and G and alanine-valine, serum albumins) occur in the hatched area Fig. 2*a*). The probability that such a distribution is a chance deviation from what is actually a random distribution is very small ($\chi^2 = 12{\cdot}3$), so that Fig. 2*a* provides strong evidence that avoidance of replacements prohibited as a result of $\alpha$ mutation is real.†

On the other hand, the distribution of 44 replacements differentiating anciently divergent proteins is in the ratio of 32 : 13 between the two areas (Fig. 2*b*), which is close to a random distribution ($\chi^2 = 1{\cdot}4$). Comparing the two distributions, one can draw the following conclusions:

---

answer can be given to the question, but some idea of the possible size of error can be obtained by considering a few cases where experimental cross-checks have been made. Thus the peptide, now known to be N-terminal, of human hemoglobin A was originally stated to have the sequence shown in parentheses in Table 1 (Ingram, 1959) and later corrected to the sequence shown below (Hunt & Ingram, 1959). The differences are by no means small. As another example, the sequence of Tobacco Mosaic Virus protein as determined in two different laboratories is also shown in Table 1 (Anderer, Uhlig, Weber & Schramm, 1960; Tsugita *et al.*, 1960). In this case it is not clear to what extent the differences may be due to the use of slightly different strains of virus, but certainly at least some must be the result of errors in one laboratory or another. As these examples are taken from the published work of established investigators with a reputation in their field, it is clear that a considerable number of errors may exist in the published sequences, introducing what one must consider as "noise" in the data. Such "noise" would, of course, tend to randomize the apparent distribution of replacements.

† If the occurrences of any replacement were equally probable, 0·2158 of all replacements should be found in the hatched area. However, the actual probability of any given replacement should be proportional to the product of the abundances of the two amino acids in the protein sample. Taking the average abundance of amino acids to be that given in Table V of my 1958 paper, the sum of such products occurring in the hatched area is 0·212, virtually identical with the number given above. Replacements forbidden as a result of $\alpha$ mutation are, therefore, not a class whose occurrence is *a priori* less probable on statistical grounds. On a random basis the probability that of 62 replacements two or less will occur in the hatched area is 0·000035.

1. Replacements differentiating the recently divergent proteins strongly avoid transitions which the coding hypothesis prohibits as a result of mutation in the $\alpha$ component.
2. Randomization of replacements differentiating the anciently divergent proteins indicates that the $\beta$ component can also mutate, but at a much slower rate than the $\alpha$ component.
3. The random distribution of replacement in Fig. 2*b* speaks strongly against the view that the non-random distribution in Fig. 2*a* is due to a selection process, which makes certain transitions unobservable because they result in an inviable organism. This suggestion is in any case implausible, since many replacements plotted in Fig. 2*a* involve amino acids of very different types, such as glutamic-lysine, or arginine-serine.

The concrete assignment of replacements into two categories was based on considerations that had nothing to do with the replacement process (Yčas, 1960), so that the observed avoidance of replacements prohibited by $\alpha$ mutation is strong independent evidence in favor of the original coding hypothesis. Since the hypothesis was deduced from data on the composition of viruses, while the data on replacements is mainly from non-viral proteins, it is further evident that a ribonucleic acid analogous in function to that of a virus must be presumed to be present in uninfected cells as part of the normal protein specifying mechanism. Identification of the $\beta$ component remains a problem for future research.

This work has been supported by Grant G-9753 from the National Science Foundation.

## REFERENCES

Acher, R. (1960). *Annu. Rev. Biochem.* **29,** 542.
Acher, R., Chauvet, J., Lenci, M. T., Morel, F. & Maetz, J. (1960). *Biochim. biophys. Acta,* **42,** 379.
Allison, A. C., Cecil, R., Charlwood, Y. A., Gratzer, W. B., Jacobs, S. & Snow, N. S. (1960). *Biochim. biophys. Acta,* **42,** 43.
Anderer, F. A., Uhlig, H., Weber, E. & Schramm, G. (1960). *Nature, Lond.* **186,** 922.
Anfinsen, C. B., Aquist, S. E. G., Cooke, J. P. & Jonsson, B. (1959). *J. biol. Chem.* **234,** 1118.
Braunitzer, G., Hilschmann, N. & Müller, R. (1960a). *Hoppe-Seyl. Z.* **318,** 284.
Braunitzer, G., Liebold, B., Müller, R. & Rudloff, V. (1960b). *Hoppe-Seyl. Z.* **320,** 170.
Braunitzer, G., Rudloff, V., Hilse, K., Liebold, B. & Müller, R. (1960c). *Hoppe-Seyl. Z.* **320,** 283.
Du Vigneaud, V., Lawler, H. C. & Popenoe, E. A. (1953a). *J. Amer. chem. Soc.* **75,** 4880.
Du Vigneaud, V., Ressler, C. & Trippett, S. (1953b). *J. biol. Chem.* **205,** 949.
Elliot, D. F. & Peart, W. J. (1956). *Nature, Lond.* **177,** 527.
Felix, K. (1952). *Experientia,* **8,** 312.
Gamow, G., Rich, A. & Yčas, M. (1956). In "Advances in Biological and Medical Physics" Vol. IV, 23. Academic Press, New York.

HILL, R. L. & SCHWARTZ, H. C. (1959). *Nature, Lond.* **184,** 641.
HIRS, C. H. W., MOORE, J. & STEIN, W. H. (1960). *J. biol. Chem.* **235,** 633.
HUNT, J. A. & INGRAM, V. M. (1959). *Nature, Lond.* **184,** 640.
HUNT, J. A. & INGRAM, V. M. (1960). *Biochim. biophys. Acta,* **42,** 409.
INGRAM, V. M. (1955). *Biochim. biophys. Acta,* **16,** 599.
INGRAM, V. M. (1959). *Brit. med. Bull.* **15,** 27.
ISHIHARA, Y., SAITO, T. & FUJINO, M. (1958). *Nature, Lond.* **181,** 1468.
KENDREW, J. C., PARISH, R. G., MARACK, J. C. & ORLANS, E. S. (1954). *Nature, Lond.* **174,** 946.
LI, C. H., PARCELLS, A. J. & PAPKOFF, N. (1958). *J. biol. Chem.* **233,** 1143.
LORAND, L. & MIDDLEBROOK, W. R. (1953). *Science,* **118,** 515.
MONIER, R. & JUTISZ, M. (1954). *Biochim. biophys. Acta,* **15,** 62.
NIU, C. I. & FRAENKEL-CONRAT, H. (1955). *Arch. Biochim. Biophys.* **59,** 538.
PARCELLS, A. J. & LI, C. H. (1958). *J. biol. Chem.* **233,** 1140.
PETERS, T., LOGAN, A. C. & SANFORD, C. A. (1958). *Biochim. biophys. Acta,* **30,** 88.
PORTER, R. R. & SANGER, F. (1949). In "Haemoglobin". (F. J. W. Roughton and J. C. Kendrew, Eds.) Interscience, New York.
SANGER, F. (1960). *Brit. med. Bull.* **16,** 183.
SKEGGS, L. T., LENTZ, K. E., KAHN, J. R., SHUMWAY, N. P. & WOODS, K. R. (1956). *J. exp. Med.* **104,** 193.
ŠORM, F., KEIL, B., HOLEYŠOVSKÝ, V., KNESSLOVÁ, V., KOSTKA, V., MÄSIAR, P., MELOUN, B., MIKEŠ, O., TOMÁŠEK, V. & VANĚČEK, J. (1957). *Coll. Trav. chim. Tchécosl.* **22,** 1310.
THOMPSON, E. O. P. (1954). *J. biol. Chem.* **208,** 565.
TSUGITA, A. & FRAENKEL-CONRAT, H. (1960). *Proc. nat. Acad. Sci. Wash.* **46,** 636.
TSUGITA, A., GISH, D. T., YOUNG, J., FRAENKEL-CONRAT, H., KNIGHT, C. A. & STANLEY, W. M. (1960). *Proc. nat. Acad. Sci. Wash.* **46,** 1463.
TUPPY, H. (1959). *Naturwissenschaften,* **46,** 35.
WALDSCHMIDT-LEITZ, E. (1953). *Angew. Chem.* **65,** 349.
WHITE, W. F. & LANDMAN, W. A. (1955). *J. Amer. chem. Soc.* **77,** 1711.
YAMAMOTO, M., KOTAKI, A., OKUYAMA, T. & SATAKE, K. (1960). *J. Biochem.* **48,** 84.
YČAS, M. (1958). "The Protein Text". *Symposium on Information Theory in Biology.* (H. P. Yockey, Ed.) Pergamon Press, New York.
YČAS, M. (1960). *Nature, Lond.* **188,** 209.

*J. Theoret. Biol.* (1961) **2**, 258–287

# Quantum Conversion in Photosynthesis†

Melvin Calvin‡

*Department of Chemistry and Lawrence Radiation Laboratory, University of California, Berkeley, California, U.S.A.*

(*Received* 31 *January* 1961)

A new suggestion is made based on model work associated with similar measurements on the biological material itself. The primary quantum conversion act is an ionization occurring in a charge transfer complex This is what it amounts to in chemical terms. But this process cannot occur in isolated charge transfer molecules in solution because the products cannot escape from each other. The primary quantum conversion as it occurs in modern photosynthesis can only take place in a laminated structure where the electrons and holes can escape from each other by electron migration and not by atomic migrations. This is the essential feature introduced here which differs from all the previous notions of how quantum conversion occurs in chemistry or biology.

## 1. Introduction

One can hardly begin a discussion of the problem of photosynthesis, or any specific aspect of it, without writing a small equation which will define and delimit the discussion. The overall reaction of photosynthesis, the reaction by which green plants convert electromagnetic into chemical energy, is usually written in this form:

$$CO_2 + H_2O \xrightarrow{h\nu} (CH_2O)_n + O_2$$

The substances on the left-hand side of the equation ($CO_2$ and $H_2O$) are the elements of carbon, oxygen and hydrogen in their lowest energy form, and the substances on the right-hand side of the equation (carbohydrate and oxygen) represent these same elements at a higher chemical potential. The carbohydrate and the oxygen normally, in the animal body and in the plant too, for that matter, can back-react, producing carbon dioxide and water and, at the same time, liberate energy in one form or another—energy for growth, energy for heat, energy for whatever purpose the organism might want it.

Certain aspects of this problem of energy conversion will not be the

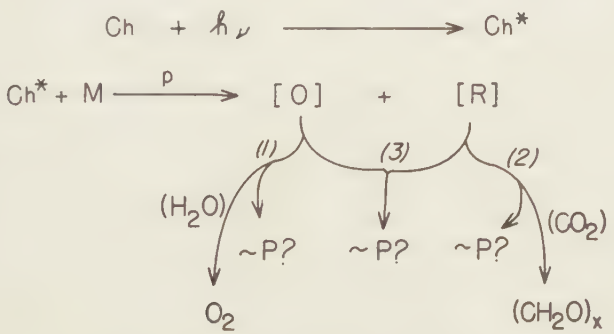
† The preparation of this paper was sponsored by the U.S. Atomic Energy Commission.

‡ Research Professor of Chemistry in the Miller Institute for Basic Research in Science, University of California, Berkeley, 1960-61.

subject of this discussion, either because they have been resolved or because we know little about them. In the first category lies that part of energy conversion in which carbon passes from carbon dioxide into carbohydrates. With the use of tracer carbon, we have been able in the past fourteen years to draw an almost complete road map from carbon dioxide to the various chemical compounds that go to make up the plant (Bassham & Calvin, 1957, 1960, in press; Bassham, 1959), principally carbohydrates. In the second category is the conversion of the oxygen from water to molecular oxygen, at the opposite end of our knowledge level; we know nothing, really, about how the single oxygen atom in the water molecule finds another one and becomes an oxygen molecule—in other words, how the oxygen-oxygen bond is created. We have some ideas about it, but very few in contrast to what we know about the construction (the actual building) of carbon compounds. We know very little about how we put together an oxygen molecule (Dorough & Calvin, 1951; Anderson, Blass & Calvin, 1959; Sapoznikov, Eidelman, Bazhanova & Popova, 1959; Mason, 1957).

In between these two phases of our knowledge of the process of photosynthesis and energy conversion lies the area of the present discussion. It is the aspect in which the electromagnetic quantum (the light quantum) is absorbed by the chlorophyll to give an excited electronic state of chlorophyll, and then something happens to this excited electronic state, during which time it is converted into chemical potential, definite molecular species which, upon back-reaction, could liberate energy. That particular step is the primary concern of this paper.

To isolate for consideration that step from the equation as it is written, we may describe the events as follows:




The quantum is first absorbed by the chlorophyll molecule; then something happens (p for primary) to the excited chlorophyll to produce two chemical species ([O] and [R,] for example) which later can go on, one of

them [O] to become molecular oxygen in some way (*1*), and the other one [R] leading to the reduction of carbon dioxide to carbohydrate (*2*). Along these two routes various other energy-containing species may be created, such as phosphoric anhydride (ATP or $\sim$P). A phosphoric anhydride species, represented by ATP, would, of course, be an energy storage product. These may be created on either, or both, sides. Further than that there may even be back-reaction (*3*) between these intermediates (oxidants and reductants) which also could create various products of higher energy. The obvious one to use here is, of course, the pyrophosphate linkage. The creation of a pyrophosphate linkage of this sort in a water milieu is storing energy.

## 2. Photochemistry of Chlorophyll

No attempt will be made to describe the biochemical detail of any of the steps beyond (p). This paper will be limited to the very first thing that happens to the quantum after it has been absorbed by the chlorophyll molecule to produce an excited state of the chlorophyll. What are the very first forms in which stable (definable) chemical species different from electronically-excited molecules (such as excited chlorophyll) appear? I shall not be concerned with how the intermediate oxidant [O] becomes oxygen (*1*) or what other intermediate oxidants might be, nor shall I consider what the hydrogen carriers might be which eventually reduce carbon dioxide to carbohydrate (*2*) or how, along the line (*2*) as they drop in potential, they might produce other high energy containing materials such as ATP. The recombination (*3*) oxidant and reductant which might also occur as succeeding chemical steps, will also lie outside our present concern. My concern is the immediate fate of the excited chlorophyll and what could possibly be the very first of these species here called oxidants and reductants.

In order to try and get some idea of what could happen to the excited chlorophyll, two additional ideas will be introduced. First of all, I shall examine the biological apparatus that performs this operation (insofar as we know what molecules that biological apparatus is made of and how it is constructed), and, secondly, I shall explore some model experiments which are based upon what we believe is the construction of this biological apparatus. (This latter is almost exclusively physical chemistry or physical-organic chemistry.) Then I shall go back and apply the concepts that are devised from the combination of the structural information and our model researches, to the biological material itself, experimental observations on the biological material designed to simulate or reproduce the observations that were made on the model systems.

## PHOTOCHEMISTRY OF CHLOROPHYLL IN SOLUTION

Before going into the details of this, it seems worthwhile to introduce the point of view that dominates these discussions. From the very beginning of our knowledge of the structure of chlorophyll (in 1911 Willstätter & Stoll (1939) first had a good idea of what the structure was), chemists, biologists and biochemists tried to understand the photochemistry of chlorophyll itself. As they extracted chlorophyll from leaves of green plants and worked on the structure of it, they studied its photochemical behavior as well. The Fischer formula has since been confirmed completely (Woodward *et al.*, 1960), and we can now proceed with complete confidence in it.

Over a period of some forty years photochemists made a wide variety of experiments in an attempt to see how the energy of a 40 kcal quantum (which is what is involved here) could be converted in a single act into chemical potential. An enormous literature (Gaffron, 1933; Schenck, 1957; Krasnovskii, 1960; Livingston, 1960) exists on the photochemistry of chlorophyll and models of it. A great many attempts have been made to find ways in which the energy of 40 kcal in an excited electronic state might be used in a single act to create two chemical species that potentially could back-react with about 40 kcal—in other words, to store almost all of that 40 kcal. Even if only 35 kcal were stored, that would be a lot to store in particles created at the same point. This search has not been successful, in spite of forty years work, and the many men's lives involved in it. The attempt to find a chemical reaction, either sensitized by chlorophyll or by any of its analogs or by model substances representing it, in which the energy of 40 kcal could be converted into a pair of chemical species storing something of the order of 30–35 kcal (the efficiency of this process must be very high) has not succeeded.

In retrospect, it is not very surprising that it should have not yet succeeded. If this energy conversion process is to take place in chlorophyll molecules which are simply moving about at random in ordinary solution, and in contact with a variety of molecules with which they could react and to which they could give energy, it is necessary to create, in one operation, a pair of energy-rich species A and B.† Then A + B by definition, in their back-reaction have 35 kcal of energy to set free, and they have to be created in one act right on or near the chlorophyll molecule. Some rather tricky kinetics must be involved. Most chemical reactions do not have activation energies that high; usually they are only around 20 kcal. To store 35 kcal from the starting point (let us define A·B as the starting point: this could

† These may be in different parts of the same molecule in which case the photoreaction might be called a rearrangement.

be a molecule or molecular system) the end product, A + B, has to be 35 kcal above it. If this product is not to return immediately, there has to be a barrier between it and the starting point so that the system will not fall back immediately in the back-reaction. This cannot be done; if we are to store 35 kcal and we have only 40 kcal in the quantum with which to do it the barrier cannot be more than 5 kcal high and the back-reaction would be too fast. Essentially the problem is: To separate the products which are themselves of high potential energy for reaction before back-reaction can take place. This is very hard to do in ordinary statistical chemical reactions. In fact, it has not yet been done.

There are a number of cases in which the photochemist has succeeded in storing energy in a straightforward photochemical reaction in solution, but, in general, those storages are very small, a few kcal at most, and 40–60 kcal quanta are used to accomplish this. The situation, therefore, is just the reverse of the natural reactions of chlorophyll. Instead of the product being 35 kcal above the starting point, it is only 5 kcal, with a 50 kcal quantum to help, and the barrier can be quite high (45 kcals by these numbers). One can succeed in that kind of storage problem.

The point of view adopted here is that this 35 kcal energy storage is *not* the result of ordinary statistical photochemistry in solution, but rather is the result of a *photophysical* process in an organized solid or quasi-solid matrix. How this is achieved in this case, in contrast to solution chemistry, will be the substance of this discussion. We did model work to show that this was possible in model systems. We then went on to ask if the phenomena we see in the model systems could be reproduced in the biological material itself.

### 3. Photophysical Effects in Model Systems

#### (A) ENERGY TRANSFER IN MODEL SYSTEMS

One of the factors that contributed to the adoption of this viewpoint was the examination of the structure of the biological apparatus that accomplished the energy conversion (Steinmann & Sjostrand, 1953; Frey-Wyssling, 1957). Plate I shows the chloroplast of a green plant in which this energy transfer occurs. The chloroplasts contain the chlorophyll, and it is in these (a few microns in size) that the energy conversion process occurs. Plate II is an electron micrograph of a single chloroplast, at much higher magnification, which shows the internal structure of one of the chloroplasts shown in Plate I. There is a very high degree of organized structure to be seen inside the chloroplasts. The dark areas are the so-called lamellae which are present in all photosynthetic organisms. In this particular one (tobacco) these lamellae are arranged in stacks, and the term "granum" has been applied to a single one of these ellipsoidal packages which can be

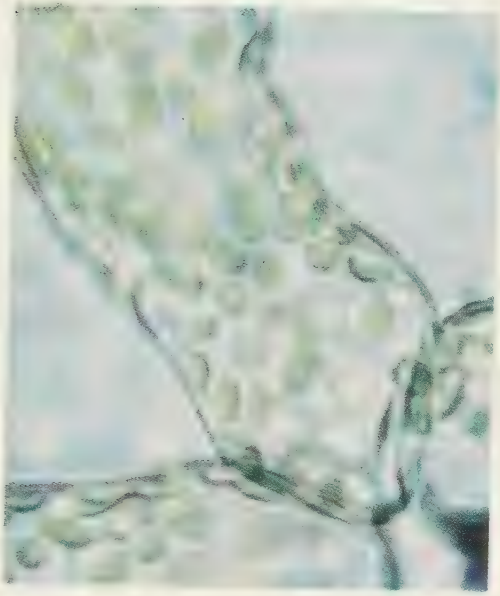


PLATE I. Cells of liverwort showing chloroplasts.

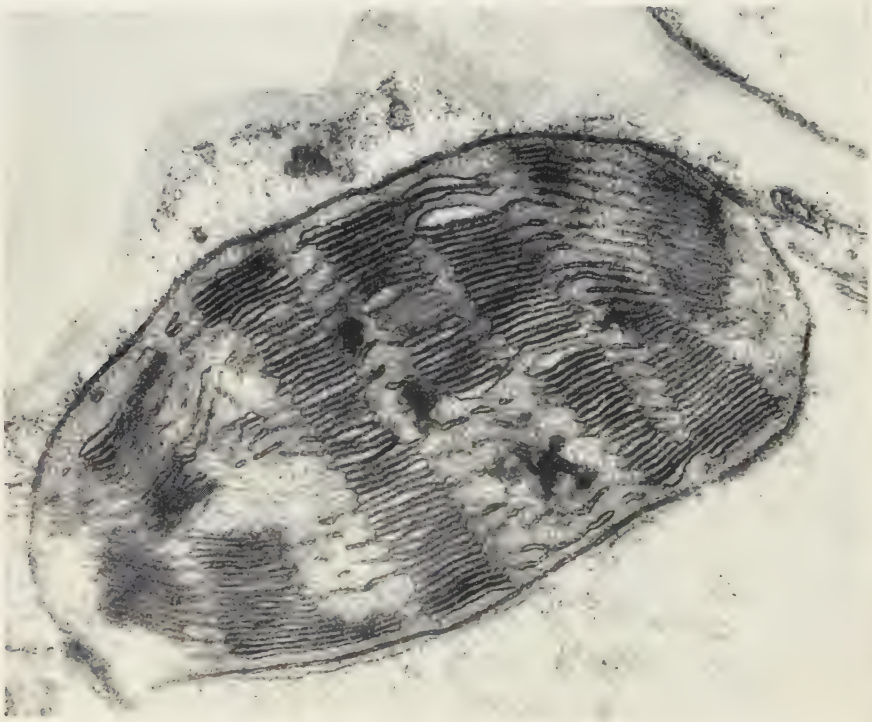


PLATE II. Tobacco Chloroplasts. 24–36 hr. in dark before fixing with permanganate (Weier, University of California, Davis).

separated from the chloroplasts. There is, then, a high degree of order to be found inside the chloroplast. In fact, if one takes a smaller section of this granum at still higher magnification, one can see that these are made up of what look like little oval sacks pressed together. The darkest areas appear to be the contact areas between the two surfaces of completely enclosed oval or ellipsoidal sacks.

Figure 1 shows a diagram of our concept of what the layers of the chloroplast are composed of (Park & Pon, 1961). Each of the dark areas represents a contact between the surface of two of the ellipsoidal packages. We believe, for the moment at least, that these are the areas in which the chlorophyll is located. We do not know as yet what the orientation of the various pigments (the chlorophyll and carotenoid pigments which are

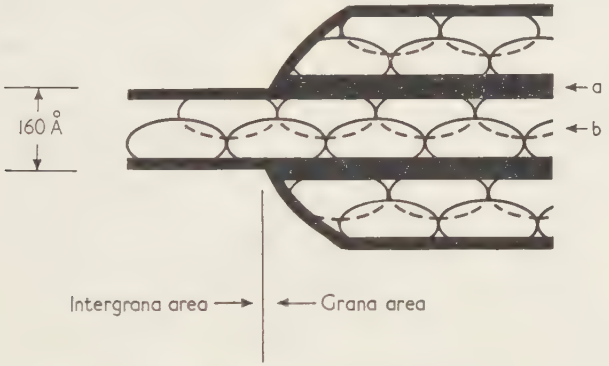



FIG. 1. Model for chloroplast lamellar structure (Park & Pon, 1961).

present in this structure) is with regard to the particular boundaries that are visible in the electron microscope. This is one of the major problems which the physical chemists working with macromolecular materials must begin to solve.

We believe that, because a high degree of order is visible, insofar as we have been able to see it, we shall find (when the research apparatus and facilities are able to carry us further) that this high degree of order is actually carried right down to the molecular level, right down to the chlorophyll plane itself and the carotenoid as well. While we could draw a picture as to what we think the molecular orientations may be, these pictures are in a large measure conjecture (Calvin, 1958). There are certain bits of evidence such as polarization, absorption and fluorescence which lead us to believe that the chlorophyll planes are oriented in these particles. The strongest argument, however, is our conviction that there must be order down to the molecular level so as to provide order at this higher level (Calvin, 1961).

There is another bit of evidence which is seen in Fig. 2. This shows the absorption spectra of chlorophyll in various states: in solution; in a colloidal state; and in a crystalline layer. The absorption spectrum of chlorophyll in the plant itself resembles the latter two more than the first one.

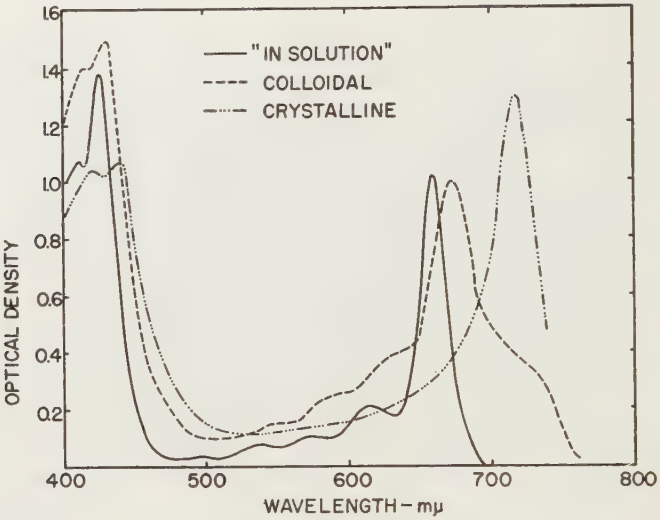



FIG. 2. Absorption spectra of chlorophyll in various states (Jacobs, Holt & Rabinowitch, 1957).

Evidently the plant chlorophyll is not chlorophyll in solution: it is lipid, protein and chlorophyll (with other pigments) in a tight package; in a semicrystalline form. I am not emphasizing the spectrum itself as the only bit of evidence, but simply as one piece indicating the ordered array which the chlorophyll in the chloroplast itself is likely to turn out to have when we know it.

### (B) RELATIONS BETWEEN CHLOROPHYLL, PROTOCHLOROPHYLL AND BACTERIOCHLOROPHYLL

What is the molecule we are talking about? Fig. 3 shows three of the chlorophylls with which we are normally concerned. The middle structure shows chlorophylls *a* and *b*; chlorophyll *a* has a methyl group in the 3-position and chlorophyll *b* has a formyl group (formaldehyde) in that position. Bacteriochlorophyll is found in all the photosynthetic bacteria which do not make oxygen but which do reduce $CO_2$. The essential difference between plant chlorophyll and bacteriochlorophyll is the fact that the latter has two extra hydrogens on the opposite pyrrole ring (at positions 3 and 4) as compared to a double bond for the plant chlorophyll;

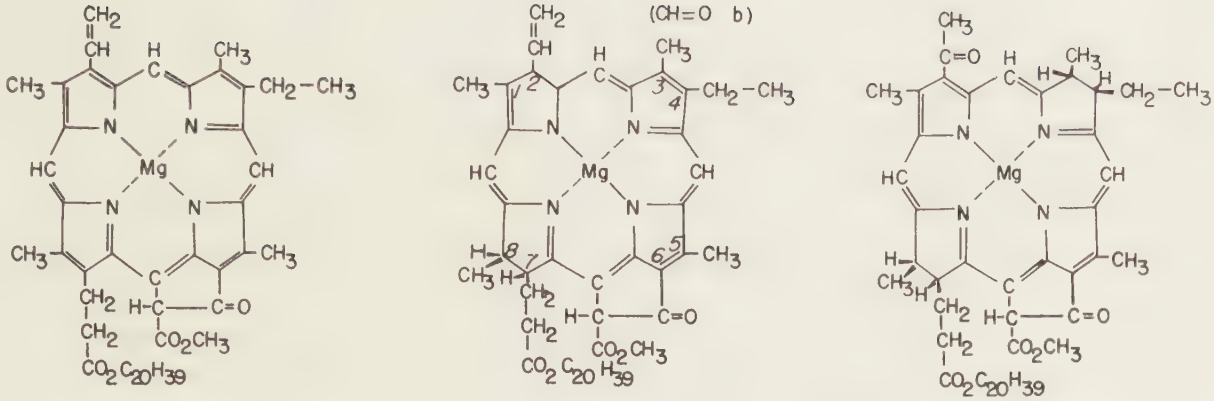



FIG. 3. Structures of protochlorophyll, chlorophyll *a* and *b* and bacteriochlorophyll.

the total redox level remains the same, since the 2-vinyl group is now oxidized to acetyl. The hydrogen atoms are just at a different place. In both the plant chlorophyll and bacteriochlorophyll, the macrocycle remains conjugated, but it is somewhat more limited in the bacteriochlorophyll.

Protochlorophyll belongs to the class of compounds known as porphyrins; it is dehydrogenated at positions 7 and 8 compared with chlorophyll, but that is the only difference between them. Protochlorophyll appears in etiolated plants, that is, plants grown in the dark from seed and which have never seen the light. Protochlorophyll is converted into chlorophyll immediately upon illumination (Smith & Coomber, 1960). These "extra" hydrogens, the 7 and 8 pair and the 3 and 4 pair, are the two points of the chlorophyll that workers (including ourselves) have focused their attention on for the last twenty years in an attempt to do solution photochemistry (Seely & Calvin, 1955). We thought that perhaps one or the other of these pairs of hydrogen atoms was being transferred back and forth by the photochemical reaction, but now the evidence seems to indicate that this is not the case and the chlorophyll does not function in this way.

The main feature of the chlorophyll structure is the big conjugate macrocycle, the so-called dihydroporphyrin ring (chlorin ring) which is the light-absorbing entity of the photosynthetic apparatus. This is what makes plants green. The phytol side chain seems to be part of the architecture that holds the molecule in place. I do not believe the phytol chain plays a part in the energy transmission, at least not directly. The 6800 Å -40 kcal quantum is absorbed by the electronic system of this conjugated macrocycle with the magnesium in the center, and from there we do not know what happens. This is what we are trying to discover and are speculating about.

Presumably, a very similar process goes on in the bacteria with the bacteriochlorophyll, the difference being that, in the bacteria, oxygen is not liberated. The primary oxidant is instead reduced by some chemical reducing agent other than water.

So much, then, for what we know about the biological equipment that is to perform the energy conversion that was described earlier. I have not mentioned the accessory pigments of which there are several and of which at least one will probably turn out to be as important as chlorophyll. People generally overlook this, although a little thought reveals that it shouldn't really be overlooked. Wherever there is chlorophyll, wherever there is photosynthesis, there is also carotenoid. In general, people have tended to ignore this, or at least have not given enough weight to the fact that somehow carotenoid and photosynthesis must be very closely associated. The carotenoid is the long conjugated carbon chain (polyisoprene

with 10 to 12 double bonds in it and some oxygen at each end), and a variety of functions have been proposed for it: oxygen carrier (Dorough & Calvin, 1951), electron carrier (Calvin, 1958; Platt, 1959), or hydrogen carrier (Calvin, 1959a; Shlyk, Godnev, Rotfarb & Lyakhovich, 1957), and probably one of them is right, but the difficulty is to know which one.

With this structural background on the photobiological apparatus, let us turn first to the question of generating an idea as to how it might work (other than ordinary solution photochemistry) in the solid state, i.e., the organized state which very certainly exists. Then I shall describe some of the model experiments that have been done in an attempt to expand, or explore, the concepts that were generated by the combination of knowing the fact that there is such a fine structure; that the flat chlorophyll molecules tend to lie one upon the other; and that there is a difference between the way the crystal or pseudo-crystal behaves and the way the molecules in solution behave.

### (c) PHTHALOCYANINE AS A MODEL FOR CHLOROPHYLL ENERGY TRANSFER

About 1950 the developments in solid state physics finally reached the chemists. By this I mean the developments in our knowledge of the electrical and magnetic properties of atomic and ionic crystals had reached a stage, both of technical development and understanding, which allowed us to apply some of the notions that were common amongst the physicists developing this work to the kinds of molecules and the kinds of systems that we had in this biological apparatus, particularly the big, flat aromatic systems such as chlorophyll.

I had for some years been working with porphyrin analogues. The first of these, and the one that is still one of the most popular, I encountered in 1936, the year it was discovered in England, and this is the molecule of phthalocyanine. It is a synthetic compound that resembles, in some respects, the structure of the tetrapyrrole visible in chlorophyll. Phthalocyanine differs from chlorophyll in certain rather important aspects, but the most important difference is that it is easily made compared with chlorophyll; it is also easily handled and very stable, and neither of these qualities is true of chlorophyll. This is the reason we selected phthalocyanine as a model of the porphyrin structure found in the chlorophyll in an attempt to find out how the solid array of molecules might differ in their physical and chemical properties and reaction to light from molecules in solution.

The structure of phthalocyanine was determined in 1935–36 by Linstead (Linstead, Eisner, Ficken & Johns, 1955) at Imperial College. It is shown in Fig. 4. It is made from phthalonitrile and metal; the ring closure occurs very readily. It has the elements of the tetrapyrrole in it, but it differs from a true tetrapyrrole in that the bridging atom instead of being

CH is nitrogen, so it is called a tetrazaporphyrin. It also has benzene rings fused on to the pyrrole rings. Phthalocyanine is a very stable substance and is widely used in various forms as a dyestuff.

With this as our starting point we sought to make systems that might resemble the laminated system that appears to exist in the chloroplast. The idea that organic substances such as phthalocyanine might be electronic conductors under certain conditions was actually born, as far as I was concerned, in a discussion with Professor Michael Polanyi (University of Manchester) at the time we received the phthalocyanine from Linstead in 1936. We did not do anything about it then except insofar as we used the phthalocyanine as a catalyst for hydrogen activation, much like platinum.

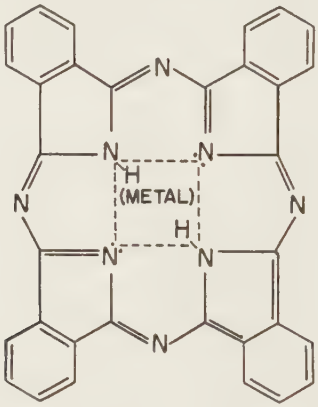



FIG. 4. Structural formula of phthalocyanine.

That was about the extent of my early activity with phthalocyanine as a possible electronic conductor (Calvin, Cockbain & Polanyi, 1936; Calvin, Eley & Polanyi, 1936). One of my associates in the laboratory at Manchester, D. D. Eley, also working with phthalocyanine, went to work along the electronic lines, and some twelve years later he published the first paper, I think, on this subject, in which he demonstrated that phthalocyanine behaved as an organic semiconductor (Eley, 1948).

This was enough to trigger us again, and now the basic idea was born that the energy conversion process in the chloroplast might be a process in which the excited chlorophyll molecule had some of the properties of an organic semiconductor. The transformation from an excited chlorophyll molecule into chemical potential was envisaged as *separation of charge rather than a separation of atoms*. We now had to devise that physical configuration of the molecules which might permit the demonstration that this phenomena could occur.

The structure of the actual photosynthetic apparatus suggests a laminated

structure in which chlorophyll molecules are arranged in some order, perhaps with carotenoids and other lipid-type materials on one side. On one side of the chlorophyll layer there could be electron-accepting species and on the other side of the layer there could be electron-donating species. In this way one could visualize a laminated system resembling the donor-acceptor systems in the atomic and ionic lattices that the physicists had been describing, which did succeed in converting electromagnetic energy into charge separation in a fairly well understood manner.

We proceeded to explore this idea and develop it to see what the limitations of it were and what the requirements were for producing charge

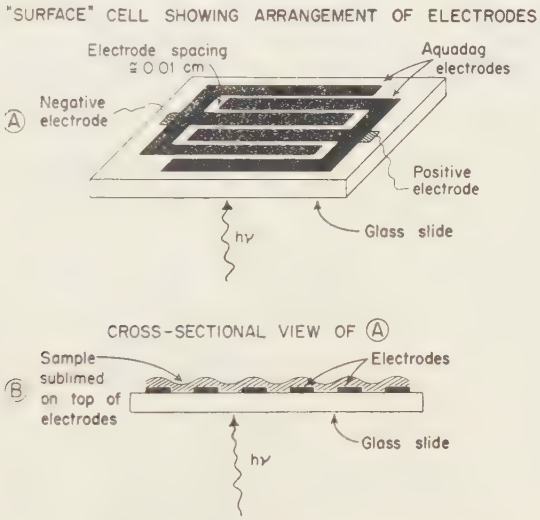



FIG. 5. Diagram of sample conductivity cells.

separation in an organic system using light. First, we had to show that the material was indeed a semiconductor. We performed the same experiments that Eley had done and came out with much the same general results. The next step was taken when we started to construct laminated (layered) structures in which we added either electron donors or electron acceptors to the phthalocyanine (chlorophyll analogue) layer (Kearns & Calvin, 1958; Kearns, 1960; Kearns, Tollin & Calvin, 1960). Our first measurements were purely of conductivity: Could these layers carry an electronic current in the dark? What would happen to the conductivity of such a system if one put donor or acceptor layers together in such a configuration?

Fig. 5 shows the diagram of the apparatus that was used to perform these experiments. The electrode system shown here was actually an interleaving of two aquadag combs, and lying on top of it, by sublimation or

evaporation, was the layer of the sample. We have performed the experiment with phthalocyanine and with about half a dozen other aromatic pi-electron-containing systems. The lamination was achieved by putting on the back surface of the sublimed layer the donor or acceptor system, whichever it might be. Most of the work on the phthalocyanine and on the other aromatic systems (violanthrene, perylene, etc.) was done with electron acceptors as the top layer (Kearns & Calvin, in press).

The results of such an experiment are shown in Fig. 6 in which we plot the log of the current flowing between the two electrodes (maintained at a

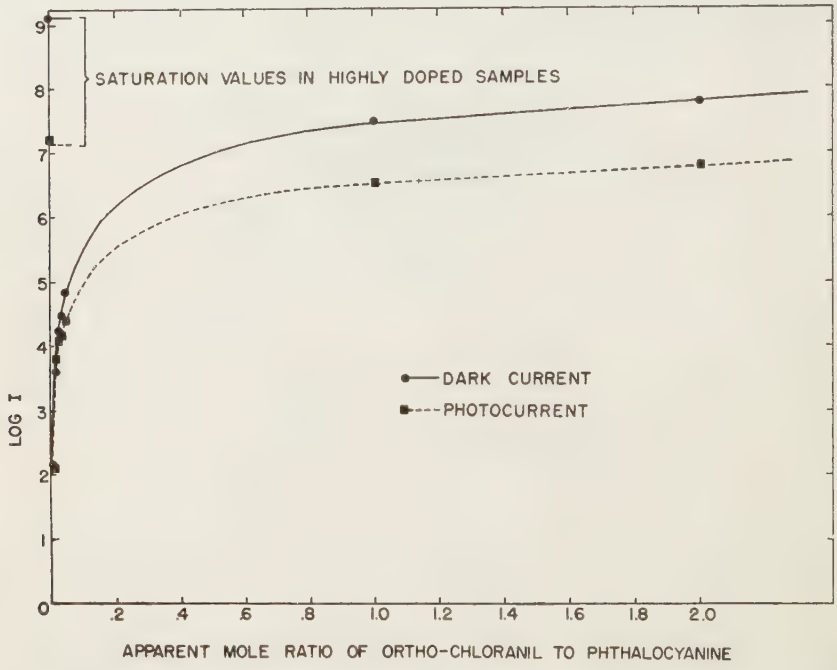



FIG. 6. Variation of dark conductivity and photoconductivity of phthalocyanine with amount of o-chloranil added.

50 to 90 V differential) as a function of the amount of electron acceptor which was put on top of the phthalocyanine layer. This, then, is the current flowing between the electrodes, i.e., through the phthalocyanine, as it is affected by the electron acceptor which is placed on top. The conductivity of this system rises very steeply as very small amounts of electron acceptor (o-chloranil) are added to the surface layer. This is true of the dark current and also of the photocurrent, which is the difference between the light current and the dark current. The o-chloranil (o-tetrachloroquinone) is a very good electron acceptor: as a small amount of the electron acceptor is

placed above the phthalocyanine layer, the conductivity goes up by several powers of ten.

Apparently the acceptor pulls electrons out of the donor, putting electrons into orbitals of the o-chloranil and leaving behind electronic vacancies in the phthalocyanine molecules. By putting a potential between the two electrodes, it becomes possible to move charge much more readily between them because there are now low-lying, *unoccupied orbitals* between which the electrons *from the full orbitals* can move. The electronic state in the organic solid after *any particular move* is the *same* as it was before, save for

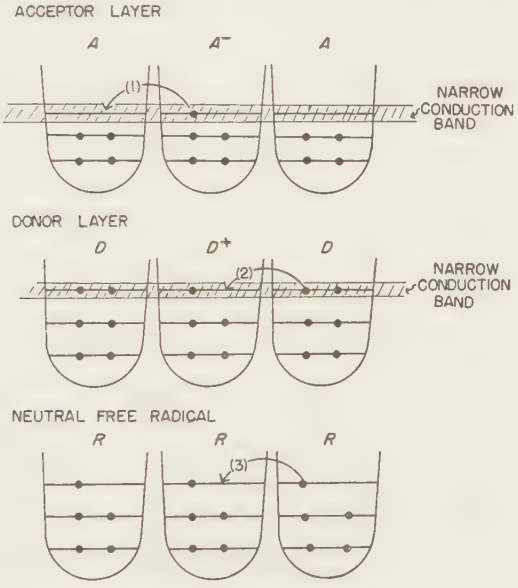



FIG. 7. Schematic representation of donor and acceptor molecules and ions imbedded in a donor layer or an acceptor layer, respectively. From this diagram it is clear that process (1), the transfer of an electron from an acceptor negative ion to a neutral neighbor, produces a state of the system which is energetically identical with the initial state. Similarly, there is no net change in energy as a result of process (2) which rearranges charge in the donor layer. In the case of a neutral free radical, however, the electron transfer process (3) does not result in a state energetically equivalent to the initial state. Since processes (1) and (2) simply change the location of negative and positive charges respectively, with no net change in energy, we can consider the orbitals involved in the electronic rearrangements as forming conduction bands. If, however, the lattice were made up of A radical ions (no A's) irrespective of the cations, or entirely of D radical ions (no D's) irrespective of the anions, there would be no identical *vacant* orbitals into which the charge carriers could move and hence no conduction bands (however narrow). This last situation would correspond to the completely filled free radical system as in process (3) above.

the passage of electrons from one electrode to the other. Without these vacancies for "hole" motion in the donor layer (electron motion in the acceptor layer), the conductivity would be very low (Keppler, Bierstedt &

Merrifield, 1960). A diagram representing this situation is shown in Fig. 7 (Kearns & Calvin, in press).

The light effect involved in the excitation of phthalocyanine to an excited state leads to a higher population of electrons in the acceptor molecules, making a higher population of electronic vacancies in the donor matrix so that the conductivity increases over that in the dark.

This is essentially the basic notion which we believe describes the model system as we now have it. We have used a wide variety of donor systems and a considerable variety of acceptor systems, and the behavior has fulfilled all of the expectations of such a description (Kearns, Tollin & Calvin, 1960; Kearns & Calvin, in press).


FIG. 8. Electron spin resonance spectrum of o-chloranil "doped" phthalocyanine. The curve represents the first derivative of absorption.

There are various other properties of such a system which should follow, and we have measured them. For example, we have measured the kinetics of the photoconductivity—how it grows and decays—at various temperatures. One observation is particularly interesting, and it has to do with the fact that in a system of this kind, the electrons in the acceptor layer are, in effect, unpaired electrons. They may be considered as in very narrow conduction bands, i.e. they are in singly occupied orbitals in the molecules. The same things may be said of the unpaired electron that remains behind. One should see those unpaired electrons by virtue of their magnetic spin resonance and indeed we have seen them in that way. Fig. 8 shows the electron spin resonance spectrum of o-chloranil "doped" phthalocyanine; the g value is very close to that of a free electron.† Fig. 9 shows the change of that signal following illumination and darkening. When the light

† g value is the gyromatic ratio defined by $\beta = \mu g \vec{H}$ where $\mu$ is induced dipole and $\vec{H}$ is the magnetic field.

is turned on, the spin signal is decreased and when the light is turned off, the spin signal comes back. The reason for that in this particular situation is that almost all of the o-chloranil molecules adjacent to the phthalocyanine are already mono-negative ions in the dark, and, when the light is turned on,

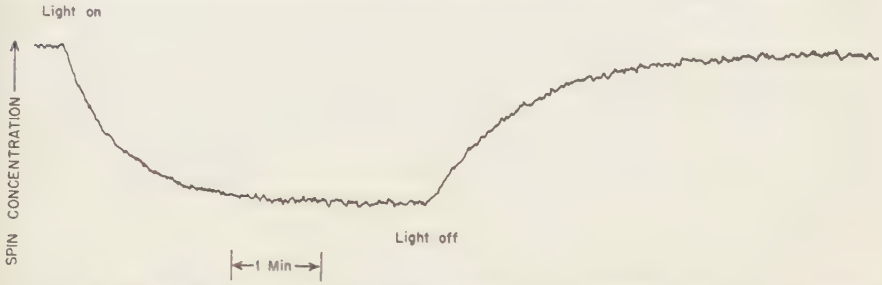



FIG. 9. Effect of illumination on the electron spin resonance signal of o-chloranil "doped" metal-free phthalocyanine. Curve represents unpaired spin concentration *v.* time.

a second electron is transferred so they become di-negative ions. Thus, there is a decrease in the total number of unpaired spins in the light. However, we do have systems that go the other way, because the equilibrium distribution is different. This depends on the relative orbital energy levels of the two systems, and we can get effects of this type ranging between photodecrease and photoincrease of unpaired spins.

Figure 10 shows how separation of charge can be accomplished in this model system if it is properly constructed. Here is a matrix of phthalo-

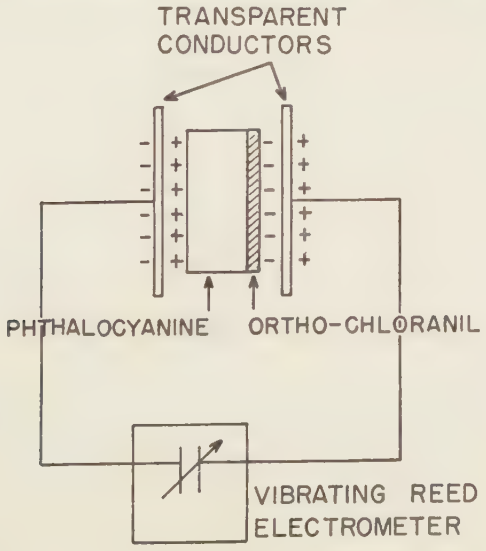



FIG. 10. Schematic diagram of polarization apparatus.

cyanine, the surface of which is an o-chloranil layer. There will be some negative charge trapped in the o-chloranil (acceptor) layer, and the positive charge will remain in the phthalocyanine (donor) layer. This will induce a polarization in the pair of electrodes between which the double layer is placed, and the polarization will be increased by shining light absorbed by phthalocyanine on the double layer, resulting in an additional accumulation of negative charge in the quinone and positive charge in the phthalocyanine. This is photochemically-induced separation of oxidizing power (positive holes) and reducing power (o-chloranil double negative ions), and presumably this kind of thing can occur in the individual layers that are seen in the chloroplasts.

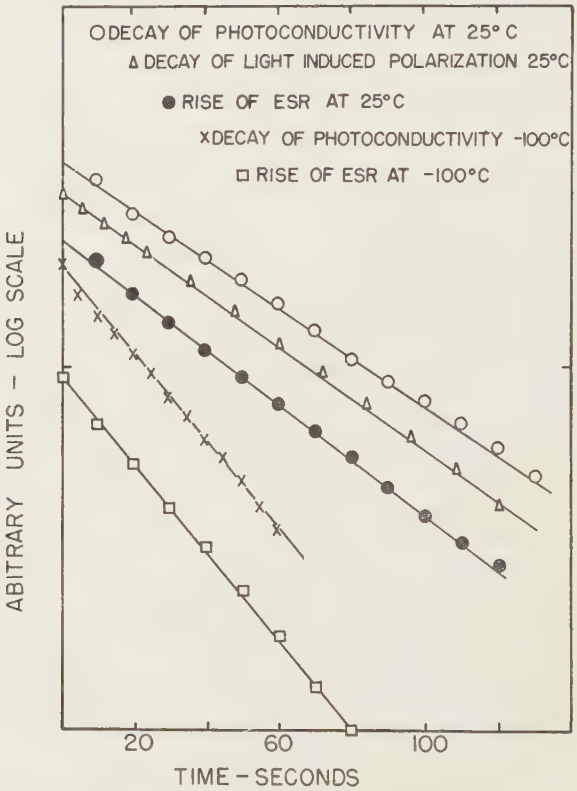



FIG. 11. Semilog plot of time dependence of photoconductivity, light-induced electron spin resonance and light-induced polarization in doped phthalocyanine.

We have studied the kinetics of various effects, the conductivity, the polarization, the electron spin resonance, and they are all apparently the result of the same process. Fig. 11 shows the kinetics of these three phenomena.

The entire system and all of the processes can be described by the series of reactions shown in Fig. 12. In the dark, the o-chloranil and phthalo-

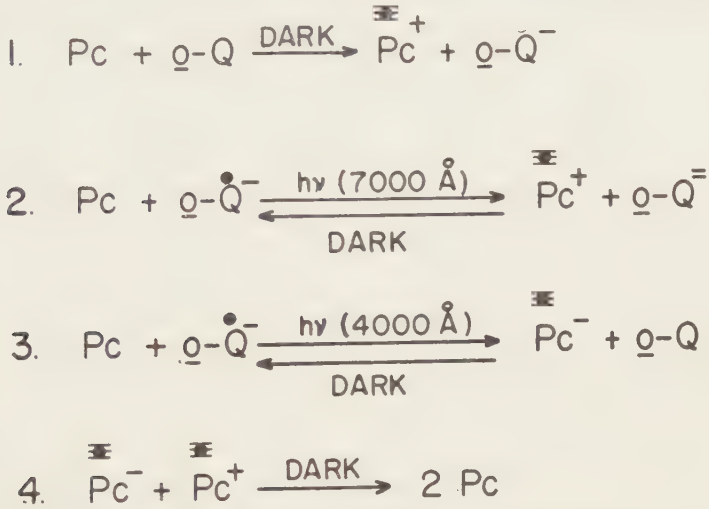



FIG. 12. Reactions of a solid matrix of Pc with a film of *o*-Q.

cyanine react to form a pair of radical ions (Fig. 12–1); in the light at 7,000 Å there is another transfer to form a double negative ion (Fig. 12–2) In the dark it goes back (Fig. 12–3). At 4,000 Å, where the semi-quinone anion absorbs, we can excite this molecule and transfer an electron back into the phthalocyanine layer, which then leads to recombination and we get a decrease in conductivity (Fig. 12–4).

This concludes what I want to say about the model experiments. I think it is clear from the model systems that organic substances of at least one type (large, aromatic molecules) can be semiconductors and photoconductors; and, what is more, by suitably adjusting the combination of donor and acceptor systems, one can make from them a laminated structure in which it is possible to demonstrate the separation of charge induced by the absorption of light, the very thing that we were postulating might occur in the chloroplasts.

### 4. The Relation to the Photosynthetic Apparatus

The remainder of the discussion is an attempt to see how many of the kinds of measurements that were performed on the model systems we can perform on the biological material, and how truly these measurements tell us what goes on in the biological material in the same manner as they tell us what goes on in the model systems.

The one thing that is difficult to do in the biological material is the very

first measurement which we made on the model system, namely, the conductivity. In the model systems we could make the configuration to fit the electrodes big enough so that we could handle it. In the biological materials, these lamina or lamella are pretty small—of the order of 30 to 60 Å thick. So far, no one has succeeded in making electrode systems that can be placed on the individual lamella to measure the conductivity or the photoconductivity of such small single units and larger ones do not seem to be available.

However, there have been conductivity measurements on dried chloroplasts which show that the dried chloroplast preparations are indeed photoconductive, but they are subject to questionable interpretation in such a complex system. Nevertheless, we shall take the measurements at their face value later on (Arnold & Clayton, 1960; Arnold & Maclay, 1958).

### (a) Electron spin resonance in chloroplast materials

One of the principal types of experiment that we have done entailed looking for the unpaired electrons that might be generated by the light in the biological system. We put the biological system inside a resonance cavity to see if any unpaired electrons were generated when the light was turned on to it. Very early we found out that there were unpaired electrons of this type. The first experiments were done with eucalyptus leaves in 1956, but we found that the results were not reproducible because of the variability of the eucalyptus leaves themselves. Towards the end of that year, the same kind of observations were made at St. Louis (Commoner, Heise & Townsend, 1956; Commoner *et al.*, 1957). We ourselves made some chloroplast preparations and investigated the same thing (Calvin & Sogo, 1957; Sogo, Pon & Calvin, 1957). This type of experiment can be done with whole organisms (whole bacteria or chromatophores which are the chloroplasts of bacteria) or with pieces of chloroplasts from the green plant.

Figure 13 shows the light-produced signals from whole spinach chloroplasts. Light of 40 kcal per quantum was directed on to these materials and there are not many chemical bonds that can be broken by as little as 40 kcal. The signal indicates the appearance of unpaired electrons. Any free radical will give this kind of signal. Most biological material that is undergoing rapid metabolism will show signals of this kind; it is not necessary to have light shining on them. The question, therefore, is: What kind of unpaired electrons are these? Are these ordinary free radicals, or are these electrons produced in photoprocesses such as have been described in the earlier models? If there were chemical free radicals produced by some secondary reactions, one might expect that if the system were cooled enough, the chemical reaction might stop and only the physical process of

electron transport would remain. We attempted to do this by cooling the sample to $-150°$ C but we still got light-induced signals.

Figure 14 shows the kinetic behavior of such a signal for *Rhodospirillum rubrum* which use bacteriochlorophyll. This experiment was done in a series of different temperatures, and it was found that the signals changed in character with the variation in temperature. There is also a variation in the signal with time. At 25° C after the light is turned on, the signal rises just as fast as the apparatus will follow it and reaches its equilibrium value

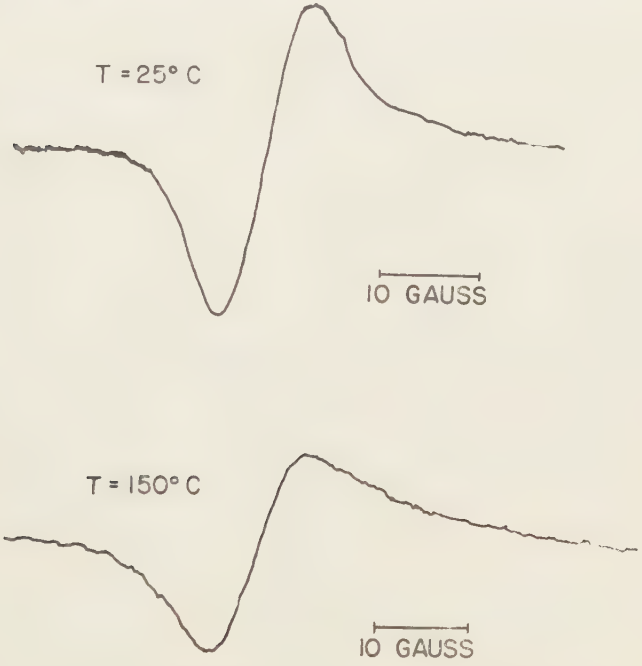



FIG. 13. Light signals from whole spinach chloroplasts.

immediately, and when the light is turned off, the signal drops as rapidly as the equipment will follow it. In other words, the rise time and decay time that we have so far been able to see are not intrinsic to the electrons but rather they are imposed by the limits of the apparatus. As the material is cooled from 25° C to $-15°$ C, a considerably larger signal appears, but there is a slow rising component in the signal. If the temperature is lowered still further to $-55°$ C, some of the "extra" signal which is purely chemical (secondary, in other words) is frozen out, but not all of it: there is still a very fast rise followed by a slow rise, and the decay time shows the same characteristic—a fast decay and a slow decay. There are quite clearly several different kinds of unpaired electrons produced in this organism

when the light is shone on it at −55° C. When the temperature reaches −160° C, we have none of the slow signals left at all, only the fast signals. Both the rise and decay are fast.

This phenomenon is most readily interpreted by the obvious notion that we are first making a conducting type of unpaired electron which then is undergoing chemical change inside the biological material, also via one-

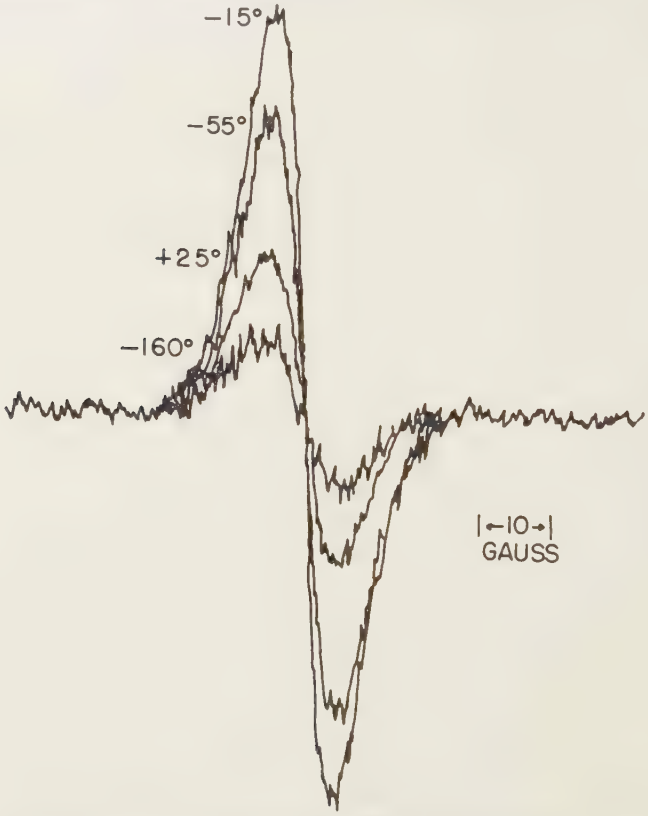



FIG. 14. Electron spin resonance signals from *Rhodospirillum rubrum*; 5 min continuous illumination.

electron reactions. We are seeing at room temperature and intermediate temperatures not only the physically-produced charge separation but chemical radicals as well, and as we cool the solution we freeze out the chemical reaction and have left only the physical process itself (Calvin, 1959b).

We really need something more to characterize the unpaired electrons. The rate of growth and decay, temperature dependence, etc., is not enough to identify these electrons as physically-produced instead of

chemically-produced. So far, the g values, i.e., the magnetic characteristics of the electron, appear to be those of free electrons, that is, electrons that are free to move around within the molecule and within the lattice.

We have tried to use one or two other ways of characterizing the electron, such as looking for hyperfine structure, that is, looking for the interaction of the unpaired electron with specific nuclei, but so far this has not been successful. Either there are so many nuclear hyperfine interactions as to overlap, or the quasi-solid matrix broadens the lines so that no very useful resolution has yet been possible (Commoner *et al.*, 1957).

### (B) APPARENT SPECTRAL EFFICIENCY

The next characterization after the kinetics and the g value was the efficiency with which light produces the spin signals—the quantum

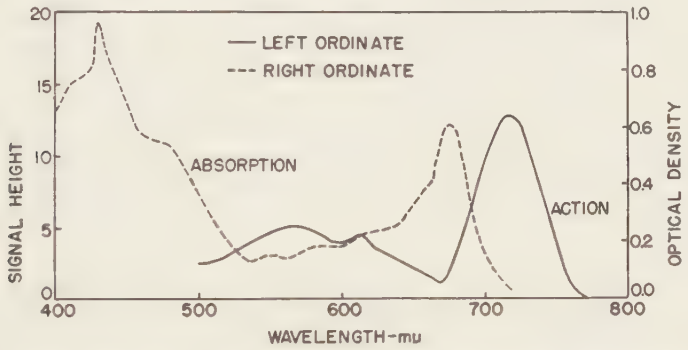



FIG. 15. Absorption and action spectra of chloroplasts (Sogo, Carter & Calvin, in press) $I_0 = 10^{15}$ quanta/sec. Band width = 100 Å.

efficiency for the production of these electrons. This is in the same vicinity as the quantum efficiency of photosynthesis, i.e., of the order of one to one-tenth. Greater precision is difficult to obtain at present.

The quantum efficiency with respect to wavelength is the next question. How does the quantum efficiency vary with wavelength? This type of measurement is somewhat easier to perform. The values that are here given are not absolute. The relative value for the production of electrons at one wavelength compared to the value for the production at another wavelength is compared with the absorption of chlorophyll. Figure 15 shows the action spectrum for the production of free electrons and the absorption spectra for the chloroplasts. It looks as though a minimum action occurs at a place where the absorption is greatest. This turns out to be what one would expect, judging from the configuration of the system. We used a thick layer of chloroplasts so that all the light was absorbed, and in those regions in which the light is most strongly absorbed, the concentration of

separated charges is the greatest and the recombination occurs at its fastest rate (Sogo, Carter & Calvin, in press). Since we are seeing production minus recombination, we see a minimum at the highest concentration of production. There is probably another effect also contributing to this shape for the "action" spectrum. It is possible to show by combinations of different wavelengths that one can get more than additive effects and less than additive effects for the sum of two or more different wavelength illuminations.

This idea of additive effects of light of varying wavelengths is a constituent part of the development of our knowledge of the behavior of plants with respect to light: it is known as the Emerson effect. In simplest terms it may be defined by the following observations; measure the number of

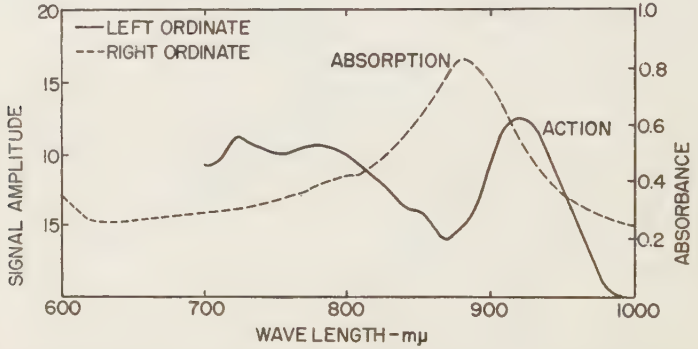



FIG. 16. Absorption spectrum (Shibata, Benson & Calvin, 1954) and action spectrum (Sogo, Carter & Calvin, in press) of *Rhodospirillum rubrum*. $I_0 = 5 \times 10^{14}$ quanta/sec. Band width = 66 Å.

molecules of oxygen produced per quantum of red light; measure the number of molecules of oxygen produced per quantum of green light; and then put both the red and the green light together on the same plant. This can be done under circumstances such that when the two wavelengths of light are together on the plant, one gets more (or less) than the sum of the two separately. In other words, there is a collaboration of the two wavelengths of light (Emerson, Chalmers & Cederstand, 1957). The experiment can be done under conditions where there is a negative collaboration of the two wavelengths (they cancel each other) depending on the light intensities and other conditions of the experiment (Govindjee, Rabinowitch & Thomas, 1960; Ichimura & Rabinowitch, 1960).

The same type of experiment can be performed with the photo-induced spin signals, at low temperatures (Androes, private communication). This is one more reason to suppose that the spin signals that we see are indeed something very close to the quantum conversion process itself.

Figure 16 shows the absorption and action spectrum for the purple

acteria, and you can see exactly the same relationship between the absorption and the action (Shibata, Benson & Calvin, 1954).

If we have achieved separation of charge in the molecular lattices and if

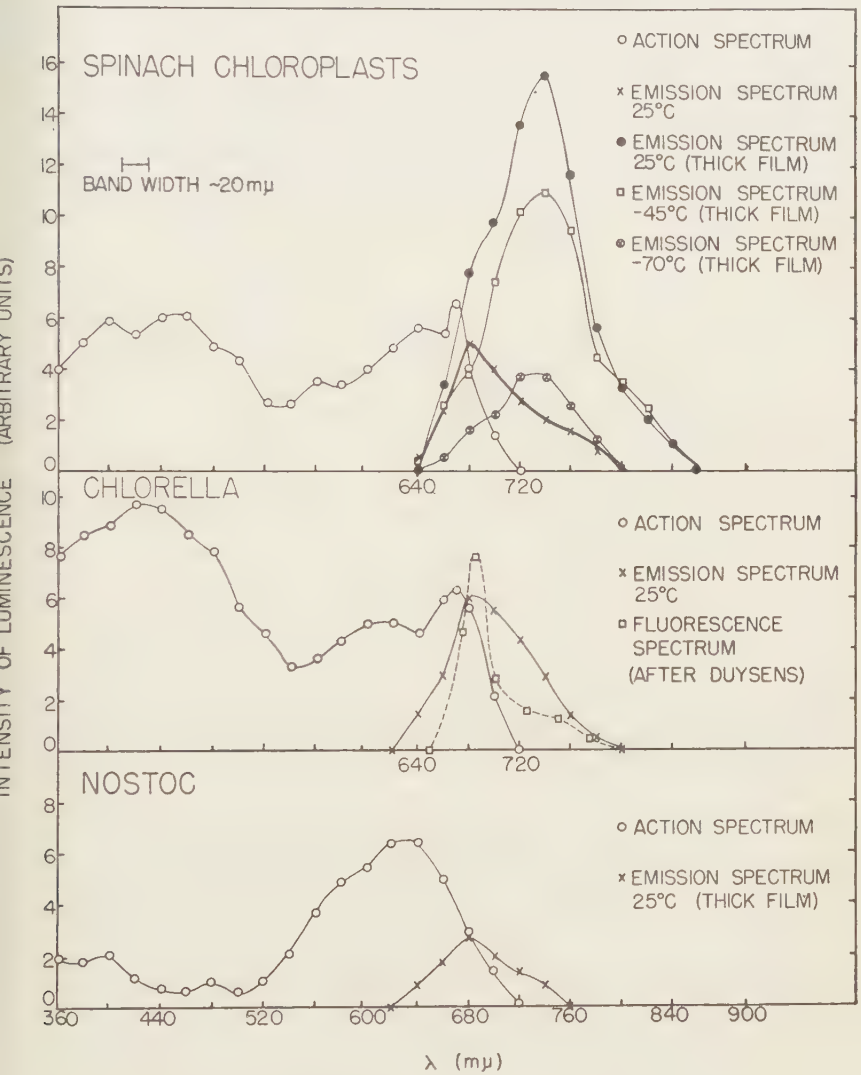



FIG. 17. Delayed light emission from different biological materials.

he charge is allowed to recombine, light can be emitted at low temperatures. Figure 17 shows the delayed light emission from spinach chloroplasts, *Chlorella pyrenoidosa*, and Nostoc. The wavelength distribution is what

one might expect, and also the kinetics of the decay of this light emission are exactly the kinetics of the decay of the spin signal (Tollin & Calvin 1957; Tollin, Fujimori & Calvin, 1958a).

Two pieces of work that have been done by W. Arnold at the Oak Ridge National Laboratory are important here (Arnold & Maclay, 1958; Arnold & Clayton, 1960). In this case, Arnold was measuring the change in the light absorption of chromatophores from *Rhodopseudomonas spheroides* (purple bacteria) induced by illumination with a second light, usually of longer wavelength. Figure 18 shows the change in absorption at 4,200 Å; the change occurs at 300° K just as fast as the instrument can measure it.

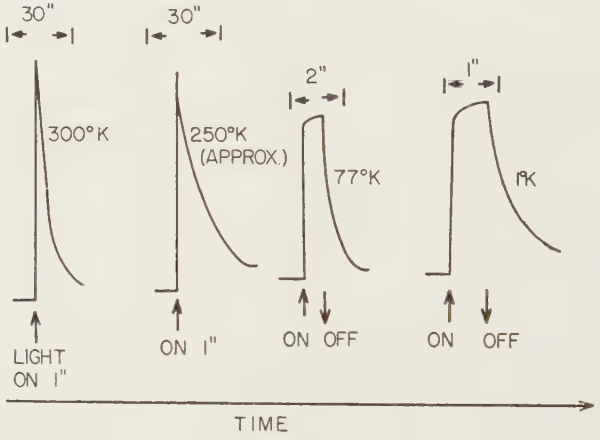



FIG. 18. Absorption changes at 4,200 Å induced by illumination with $\lambda > 6{,}400$ Å (Arnold & Clayton, 1960).

It decays relatively slowly because part of the decay is chemical and part is physical. I want to call attention to the part taken at 1° K (Fig. 18). At that low temperature, there is very little chemistry going on, and the spectral change occurs just as fast as the instrument permits measurement, in fact, faster than the instrument will follow. Here is clear evidence that the light is introducing a physical change, a change that can only be motion of electrons and not of atoms. Figure 19 shows Arnold's measurement of the photoconductivity of dried chromatophore film; again when the light is turned on, the conductivity increases very abruptly and then there is a very rapid drop.

I hope that soon we shall be able to make conductivity measurements in the radiofrequency range which do not require direct electrode connections.

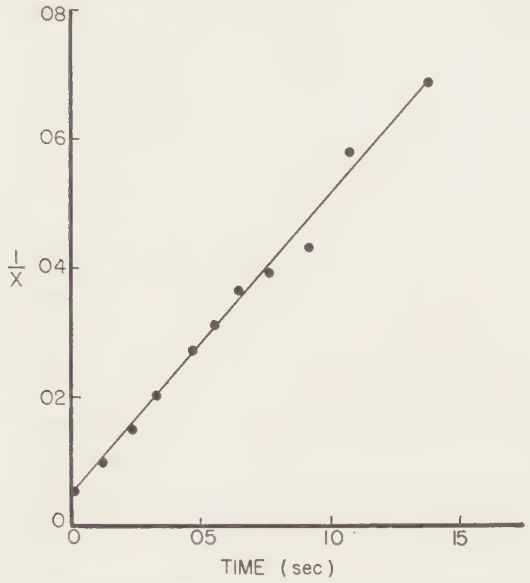


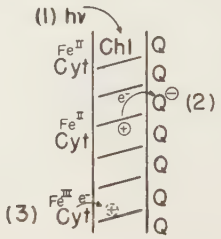

FIG. 19. Decay of transmission (X) at 6,100 Å, $T = 77°K$, following illumination with $\lambda > 6{,}400$ Å of chromatophores from *R. spheroides* (Arnold & Clayton, 1960).

### (c) QUANTUM CONVERSION IN BIOLOGICAL MATERIAL

I want to draw a picture of what I think is the primary quantum conversion process that goes on in that layer of chlorophyll, and other pigment, in the lattice. We know a little about the chemical composition of the chloroplast itself: it is a lipoprotein together with pigments. There are a number of specific molecules which are present in the chloroplast, two of them being chlorophyll and carotenoid. There are two other rather important molecules which are present in large amounts in the chloroplast and which have an important bearing on energy conversion. These systems require not only the presence of the absorber but the presence of an acceptor molecule for electron transfer to occur, and to finish this process we must have something present as a donor molecule. The other two species that are well established in the chloroplasts are (1) a very important quinone called plastoquinone (Bishop, 1959; Crane, 1959; Lester & Crane, 1959; Crane, Ehrlich & Kegel, 1960), and (2) a variety of molecules which might be donors. There is one particular type of the latter which I would like to select as a very likely donor molecule, namely, the iron heme (cytochrome) species which are always present in the chloroplasts and chromatophores (Kamen, 1956).

Figure 20 shows the two boundaries of the pigment layer. Chl are the

chlorophyll molecules in some array, possibly including carotenoids. The first act of photosynthesis is, of course, the absorption of the quantum by the chlorophyll molecule to produce an excited chlorophyll molecule. If this were a perfect atomic or ionic lattice, there would be absorption by




Cyt – CYTOCHROME AND/OR OTHER ELECTRON DONOR SYSTEMS (AQUEOUS PHASE)

Q – PLASTOQUINONE AND/OR OTHER ELECTRON ACCEPTOR SYSTEMS (TPN, LIPOIC ACID, ETC.) LIPID PHASE

Chl – CHLOROPHYLL

1. $Chl + h\nu \longrightarrow Chl^{*}$

2. $Chl^{*} + Q \longrightarrow \overset{\bullet}{Q}^{-} + \overset{\ddagger}{Chl}^{+}$

3. $\overset{\ddagger}{Chl}^{+} + Fe^{II} \longrightarrow Fe^{III} + Chl$

FIG. 20. Schematic arrangement of chlorophyll and possible donor and acceptor molecules in the chloroplast.

The system in the chloroplast might structurally bear some resemblance to the model shown here, the chlorophyll having associated with it on the one side the electron acceptor, plastoquinone, in a lipid environment, and on the other side electron donor materials, such as the cytochromes, in an aqueous environment. Following the absorption of a quantum in chlorophyll (eq. 1) it will migrate by resonance transfer to a suitable site near the quinone where electron transfer to the quinone will take place (eq. 2). The resulting vacancy can migrate by hole diffusion, that is, electron transfer from normal chlorophyll, into the vacant orbital of the neighboring chlorophyll positive ion. This process is the one that most nearly resembles the properties of a semiconductor and it permits the oxidant (chlorophyll positive ion) to separate from the reductant (electrons in the quinone orbitals) by a very nearly temperature-independent process. The oxidant then captures an electron from a suitable reducing agent, such as ferrocytochrome, thus producing a ferricytochrome and regenerating normal chlorophyll (eq. 3).

the entire lattice, but this is not so. It is a molecular lattice, in which interactions between molecules are relatively small compared with the interactions between atoms in germanium or ions in cadmium sulfide. The result is that the migration of this exciton occurs by resonance transfer between neighbouring chlorophyll molecules until it arrives at one that

is bound, or adjacent, to an electron acceptor such as quinone. The quinone of which I am speaking, i.e. plastoquinone, is one which was found in the chloroplasts as early as 1955 by Kofler (Kofler *et al.*, 1959) and it has since been shown to be uniquely characteristic of the chloroplasts and not of other parts of the plant or cell. The plastoquinone is closely related to a similar quinone known as ubiquinone which is found in the (nonphotosynthetic) mitochondria of plants and animals (Morton, 1958; Laidman, Morton, Paterson & Pennock, 1960).

Let us use the quinone as a likely electron acceptor; there is one plastoquinone molecule present for about 400 chlorophyll molecules. When the exciton reaches the chlorophyll molecule which is bound by a charge transfer complex to the quinone, ionization occurs, the electron is transferred, leaving behind in this chlorophyll molecule an electronic vacancy, or "hole". At this point, we must introduce the idea of charge migration (see legend to Fig. 7). Up until now, energy migration has been by resonance transfer of an exciton. After ionization occurs, I suggest (require, in fact) that there is a migration by an electron from a neighbouring chlorophyll molecule to the "hole", so the "hole" moves down to the next chlorophyll molecule until it comes adjacent to a ferro-heme (cytochrome). When the "hole" reaches this point, electron transfer occurs from the iron (Chance & Nishimura, 1960; Arnold & Clayton, 1960) or another donor to neutralize it, and the pigment layer is returned to its original condition.†

A separation of charge has been achieved, and the oxidized donor becomes an oxidant and the electron in the quinone is the reductant. The reductant can go on to reduce carbon dioxide (reaction (*2*), p. 4) and the oxidant can go on to generate oxygen (reaction (*1*), p. 4). ATP is required to help on the reduction of $CO_2$ and for many other energy-requiring operations. One possibility is that ATP may be generated during the passage of oxidant to oxygen (reaction (*1*), p. 4). ATP may also be generated on the reduction side (reaction (*2*), p. 4) and by recombination (reaction (*3*), p. 4) as well (see also the legend to Fig. 20).

## 5. Conclusion

What is the primary quantum conversion act? The primary quantum conversion act is an ionization occurring in a charge transfer complex. This is what it amounts to in chemical terms. But this cannot occur in

† The alternative step of electron transfer from donor to excited chlorophyll has also been suggested as a primary step (Kamen, 1961). A discussion of these two alternatives has led us (Calvin, 1961) to the present suggestion as the more likely of the two. The existence of two different primary quantum acts (Emerson, Chalmers & Cederstrand, 1957; Govindjee, Rabinowitch and Thomas, 1960; Haxo, 1960; French, 1961; Allen, Piette and Murchio, 1961) makes it not unlikely that both of these two alternatives may occur, although either one alone would be sufficient to support the whole process.

isolated charge transfer molecules in solution because the products cannot escape from each other. The primary quantum conversion act as it occurs in modern photosynthesis can only take place in a laminated structure where the electrons and holes can escape from each other by electron migration and not by atomic migrations. This is the essential feature introduced here which differs from all the previous notions of how quantum conversion occurs in chemistry or biology.

## REFERENCES


Allen, M. B., Piette, L. H. & Murchio, J. C. (1961). *Biochem. biophys. Res. Comm.* In press.

Anderson, J. M., Blass, U. & Calvin, M. (1959). *Plant Physiol.* **34,** 329.

Arnold, W. & Clayton, R. K. (1960). *Proc. nat. Acad. Sci. Wash.* **46,** 769.

Arnold, W. & Maclay, H. K. (1958). Brookhaven National Laboratory Symposia, **11,** 1.

Bassham, J. A. (1959). *J. chem. Educ.* **36,** 548.

Bassham, J. A. *J. chem. Educ.* In press.

Bassham, J. A. & Calvin, M. (1957). "The Path of Carbon in Photosynthesis". Prentice-Hall, Englewood Cliffs, New Jersey.

Bassham, J. A. & Calvin, M. (1960). In "Handbuch der Pflanzenphysiologie", Vol. V, Part 1, p. 884. Springer-Verlag, Heidelberg.

Bassham, J. A. & Calvin, M. In "Biogenesis of Natural Products ". (Marshall Gates, Ed.) Interscience, New York. In press.

Bishop, N. I. (1959). *Proc. nat. Acad. Sci. Wash.* **45,** 1696.

Calvin, M. (1958). Brookhaven National Laboratory Symposia, **11,** 160.

Calvin, M. (1959a). *Rev. mod. Phys.* **31,** 147.

Calvin, M. (1959b). *Rev. mod. Phys.* **31,** 157.

Calvin, M. (1961). In "Light and Life", 317. Johns Hopkins University Press, Baltimore.

Calvin, M., Cockbain, E. C. & Polanyi, M. (1936). *Trans. Faraday Soc.* **23,** 1436.

Calvin, M., Eley, D. D. & Polanyi, M. (1936). *Trans. Faraday Soc.* **23,** 1433.

Calvin, M. & Sogo, P. B. (1957). *Science,* **125,** 499.

Chance, B. & Nishimura, M. (1960). *Proc. nat. Acad. Sci. Wash.* **46,** 19.

Commoner, B., Heise, J. J. & Townsend, J. (1956). *Proc. nat. Acad. Sci. Wash.* **42,** 710.

Commoner, B., Heise, J. J., Lippincott, B. B., Norbert, R. E., Passoneau, J. V. & Townsend, J. (1957). *Science,* **126,** 57.

Crane, F. C. (1959). *Plant Physiol.* **34,** 128.

Crane, F. L., Ehrlich, B. & Kegel, L. P. (1960). *Biochem. Biophys. Res. Comm.* **3,** 37.

Dorough, G. D. & Calvin, M. (1951). *J. Amer. chem. Soc.* **73,** 2362.

Eley, D. D. (1948). *Nature, Lond.* **162,** 819.

Emerson, R., Chalmers, R. & Cederstand, C. (1957). *Proc. nat. Acad. Sci. Wash.* **43,** 135.

French, C. S. (1961). In "Light and Life", 471. Johns Hopkins University Press, Baltimore.

Frey-Wyssling, A. (1957). "Macromolecules in Cell Structure", Chapter 4. Harvard University Press, Cambridge, Massachusetts.

Gaffron, H. (1933). *Biochem. Z.* **264,** 251.

Govindjee, Rabinowitch, E., & Thomas, J. B. (1960). *Biophys. J.* **1,** 91.

Haxo, F. T. (1960). In "Comparative Biochemistry of Photoreactive Systems", 339. Academic Press, New York and London.

Ichimura, S. & Rabinowitch, E. (1960). *Biophys. J.* **1,** 99.

Jacobs, E. E., Holt, A. S. & Rabinowitch, E. (1957). *Arch. Biochem. Biophys.* **72,** 495.

Kamen, M. D. (1956). In "Enzymes: Units of Biological Structure and Function", 483. Academic Press, New York.

KAMEN, M. D. (1961). In "Light and Life", 483. Johns Hopkins University Press, Baltimore.
KEARNS, D. R. Thesis, University of California, Berkeley, June 1960.
KEARNS, D. R. & CALVIN, M. (1958). *J. chem. Phys.* **29,** 950.
KEARNS, D. R. & CALVIN, M. *J. Amer. chem. Soc.* In press.
KEARNS, D. R., TOLLIN, G. & CALVIN, M. (1960). *J. chem. Phys.* **32,** 1013 and 1020.
KEPPLER, R. E., BIERSTEDT, P. E. & MERRIFIELD, R. E. (1960). *Phys. Rev. Letters,* **5,** 503.
KOFLER, M., LANGEMANN, A., RUEGG, R., GLOOR, U., SCHWIETER, U., WURSCH, J., WISS, O. & ISLER, O. (1959). *Helv. chim. acta,* **42,** 2252.
KRASNOVSKII, A. A. (1960). *Ann. Rev. Pl. Physiol.* **11,** 363.
LAIDMAN, D. L., MORTON, R. K., PATERSON, J. Y. F. & PENNOCK, J. F. (1960). *Biochem. J.* **74,** 541.
LESTER, R. L. & CRANE, F. L. (1959). *J. biol. Chem.* **234,** 2169.
LINSTEAD, R. P., EISNER, U., FICKEN, G. E. & JOHNS, R. B. (1955). *Chem. Soc. Spec. Publ.* **3,** 83.
LIVINGSTON, R. C. (1960). *Radiation Research,* Suppl. **2,** 196.
MASON, H. S. (1957). *Science,* **125,** 1185.
MORTON, R. K. (1958). *Nature, Lond.* **182,** 1764.
PARK, R. B. & PON, N. G. (1961). *J. mol. Biol.* **3,** 1.
PLATT, J. R. (1959). *Science,* **129,** 372.
SAFOZNIKOV, D. I., EIDELMAN, Z. M., BAZHANOVA, N. U. & POPOVA, O. F. (1959). *C. R. Acad. Sci. U.R.S.S.* **127,** 219.
SCHENCK, G. O. (1957). *Angew. Chem.* **69,** 579.
SEELY, G. R. & CALVIN, M. (1955). *J. chem. Phys.* **23,** 1068.
SHIBATA, K., BENSON, A. A. & CALVIN, M. (1954). *Biochim. biophys. Acta,* **15,** 461.
SHLYK, A. A., GODNEV, T. N., ROTFARB, R. M. & LYAKHOVICH, YA. P. (1957). *C. R. Acad. Sci. U.R.S.S.* Biochemistry Section (English trans.), **113,** 103.
SMITH, J. H. C. & COOMBER, J. (1960). In *Yearb. Carneg. Instn.* **58,** 331.
SOGO, P. B., CARTER, L. A. & CALVIN, M. In "Symposium on Free Radicals in Biological Systems", Academic Press, New York. In press.
SOGO, P. B., PON, N. G. & CALVIN, M. (1957). *Proc. nat. Acad. Sci. Wash.* **43,** 387.
STEINMANN, E. & SJOSTRAND, F. J. (1953). *Exp. cell. Res.* **8,** 15.
TOLLIN, G. & CALVIN, M. (1957). *Proc. nat. Acad. Sci. Wash.* **43,** 897.
TOLLIN, G., FUJIMORI, E. & CALVIN, M. (1958a). *Nature, Lond.* **181,** 1266.
TOLLIN, G., FUJIMORI, E. & CALVIN, M. (1958b). *Proc. nat. Acad. Sci. Wash.* **44,** 1035.
WILLSTÄTTER, H. & STOLL, A. (1939). "Untersuchungen uber Chlorophyll". Springer, Berlin.
WOODWARD, R. B., AYER, W. A., BEATON, J. M., BICKELHAUPT, F., BONNETT, R., BUCHSCHACHER, P., CLOSS, G. L., DUTLER, H., HANNAH, J., HAUCH, F. P., ITO, S., LANGEMANN, A., LEGOFF, E., LEIMGRUBER, W., LWOWSKI, W., SAUER, J., VALENTA, Z. & VOLZ, H. (1960). *J. Amer. chem. Soc.* **82,** 3800.

## Instructions to Authors

1. The *Journal of Theoretical Biology* is intended to be a central publishing medium for theoretical work in all fields of biology.

Appropriate subjects for papers will include:

(*a*) generalized theories;
(*b*) theories of specific processes or phenomena, e.g. crossing over, learning, active transport, photosynthesis, carcinogenesis;
(*c*) theoretical discussion of specific projects, e.g. selective mutation, drug design, electronic models;
(*d*) theoretical discussion of methods, e.g. cytochemical methods and instrumentation, amplification techniques;

Papers concerned mainly with experimental work will not normally be accepted, but a paper will not be rejected simply because it contains some new experimental material.

2. Reviews will be accepted only if they include a substantial new theoretical discussion. Letters to the editor, intended to direct attention to a particular theoretical point, will be accepted.

3. *Submission of Manuscripts.* Manuscripts should be sent to Professor J. F. Danielli, Zoology Department, King's College, London, or to any member of the Editorial Board. The submission of manuscript will be taken to imply that the material is original.

Papers may be submitted in English, French or German. It is emphasized however that papers in English will reach a much larger group of readers than those in other languages. A paper submitted in French or German should have an abstract in English.

Publication will be accelerated if two copies of a manuscript are submitted.

4. *Manuscripts.* If the Journal is to serve its intended purpose, it is essential that the main body of its contents should be intelligible to all biologists. To this end, wherever possible mathematical development should be confined to appendices, and specialized terms should be clearly defined in the introduction to a paper.

*Summary.* Every paper should be preceded by a summary: this should not normally exceed 300 words, and should be intelligible to the general reader without reference to the main text. Abbreviations should not be used.

*Lay-out.* Manuscripts should be typewritten in double spacing throughout (including summary, footnotes, tables and legends) on one side of quarto paper, with wide margins. Footnotes, tables and legends for illustrations should be typed separately at the end of the manuscript.

*Running Headline.* A short running title of not more than 40 characters (including spaces), suitable for page headings, should be given if the full title is longer than this.

5. *Conventions.* The conventions used in current issues of the Journal for headings, references, etc., should be followed in preparing manuscripts. References should be listed alphabetically at the end of the paper, and abbreviated according to the *World List of Scientific Periodicals.* Citations in the text should read thus: Smith, Brown & Robinson (1957), or (Smith, Brown & Robinson, 1957). When a citation has more than two authors, the names of all of them should be given at the first citation (except for more than five authors), and on subsequent citations thus:

(Smith *et al.*, 1957). The conventions (Brown, 1957a), (Brown, 1957b) should be used when more than one paper by the same author(s) has appeared in one year. Volume number should be indicated in the typescript thus: 53.

Footnotes should be avoided whenever possible. Customary abbreviations will be used; authors may refer to *Chemical Abstracts*, to the *Biochemical Journal*, **66**, 8 (1957), and to *Notes on the preparation of papers to be communicated to the Royal Society* (1951). Certain non-standard abbreviations of chemical substances are convenient, and those listed in the *Journal of Biological Chemistry* may be used without further definition (e.g. DNA, ATP). Any other abbreviations of this kind should be used sparingly, and must be defined in a footnote at the point in each manuscript where the first such abbreviation occurs.

6. *Illustrations*. It is often convenient to make a drawing two or three times larger than it will appear in the Journal, but very large drawings should be avoided and it will be helpful if drawings and illustrations are submitted on quarto sheets. Drawings should be made in Indian ink on tracing linen, Bristol board, or faintly blue-lined graph paper; the originals should always be submitted with the manuscript, since photo-prints of drawings cannot be reproduced satisfactorily.

Line drawings will be referred to as Fig. 1, Fig. 2, etc., and photographs as Plate I, Plate II, etc. The name(s) of the author(s) and the number should be written on all illustrations, together with the indication "Top" if necessary. Besides the original illustrations *one* duplicate set must accompany the second copy of the manuscript; these may be photo-prints, and need not be of the same quality as the originals but must be clear and legible. If oversize original drawings are submitted (bigger than 8 by 10 inches) they must be accompanied by *two* sets of small photo-prints.

7. *Reprints*. Fifty reprints of each paper will be supplied free of charge; additional copies may be purchased. Order forms will be sent with proofs.

# Encyclopaedic Reference Works

## THE ALKALOIDS
**Chemistry and Physiology**
*Edited by* R. H. F. MANSKE
In 7 volumes

## THE BACTERIA
**A Treatise on Structure and Function**
*Edited by* I. C. GUNSALUS
*and* ROGER Y. STANIER
In 5 volumes

## THE CARBOHYDRATES
**Chemistry, Biochemistry, Physiology**
*Edited by* WARD PIGMAN
In 1 volume

## THE CELL
**Biochemistry, Physiology, Morphology**
*Edited by* JEAN BRACHET
*and* ALFRED E. MIRSKY
In 5 volumes

## THE ENZYMES
Second edition, completely revised
*Edited by* PAUL D. BOYER, HENRY LARDY
*and* KARL MYRBACK
In 7 volumes

## THE HORMONES
**Physiology, Chemistry and Applications**
*Edited by* GREGORY PINCUS
*and* KENNETH V. THIMANN
In 3 volumes

## THE NUCLEIC ACIDS
**Chemistry and Biology**
*Edited by* ERWIN CHARGAFF
*and* J. N. DAVIDSON
In 3 volumes

## THE PLASMA PROTEINS
*Edited by* F. W. PUTNAM
In 2 volumes

## THE VIRUSES
**Biochemical, Biological, and Biophysical Properties**
*Edited by* F. M. BURNET
*and* W. M. STANLEY
In 3 volumes

## THE VITAMINS
**Chemistry, Physiology, Pathology**
*Edited by* W. H. SEBRELL, JR.
*and* ROBERT S. HARRIS
In 3 volumes

## PLANT PATHOLOGY
**An Advanced Treatise**
*Edited by* J. G. HORSFALL
*and* A. E. DIMOND
In 3 volumes

## PLANT PHYSIOLOGY
**A Treatise**
*Edited by* F. C. STEWARD
In 6 volumes

## COMPARATIVE BIOCHEMISTRY
**A Comprehensive Treatise**
*Edited by* MARCEL FLORKIN *and* HOWARD S. MASON
In 6 volumes

*Descriptive literature available upon request*

**ACADEMIC PRESS** *New York and London*

**111 Fifth Avenue, New York 3, New York. 17 Old Queen Street, London, S.W.1.**

# International Review of CYTOLOGY

*Prepared under the Auspices of The International Society for Cell Biology*

*Edited by* G. H. BOURNE, *Emory University, Atlanta, Georgia*
*and* J. F. DANIELLI, *King's College, London*

**VOLUME 9** *1960, 424 pp., illus., $13.00 93s.*

**The Influence of Cultural Conditions on Bacterial Cytology**
*By* J. F. WILKINSON *and* P. J. DUGUID

**Organizational Patterns within Chromosomes**
*By* BERWIND P. KAUFMANN, HELEN GAY, *and* MARGARET R. MCDONALD

**Enzymic Processes in Cells**
*By* JAY BOYD BEST

**The Adhesion of Cells**
*By* LEONARD WEISS

**Physiological and Pathological Changes in Mitochondrial Morphology**
*By* CH. ROUILLER

**The Study of Drug Effects at the Cytological Level**
*By* G. B. WILSON

**Histochemistry of Lipids in Oogenesis**
*By* VISHWA NATH

**Cyto-Embryology of Echinoderms and Amphibia**
*By* KATSUMA DAN

**The Cytochemistry of Nonenzyme Proteins**
*By* RONALD R. COWDEN

AUTHOR INDEX—SUBJECT INDEX

**VOLUME 10** *1960, 409 pp., illus., $13.00 93s.*

**The Chemistry of Schiff's Reagent**
*By* FREDERICK H. KASTEN

**Spontaneous and Chemically Induced Chromosome Breaks**
*By* ARUN KUMAR SHARMA *and* ARCHANA SHARMA

**The Ultrastructure of the Nucleus and Nucleocytoplasmic Relations**
*By* SAUL WISCHNITZER

**The Mechanics and Mechanism of Cleavage**
*By* LEWIS WOLPERT

**The Growth of the Liver with Special Reference to Mammals**
*By* F. DOLJANSKI

**Cytological Studies on the Affinity of the Carcinogenic Azo Dyes for Cytoplasmic Components**
*By* YOSHIMI NAGATANI

**Epidermal Cells in Culture**
*By* A. GEDEON MATOLTSY

AUTHOR INDEX—SUBJECT INDEX

CUMULATIVE SUBJECT INDEX TO VOLUMES 1–9

**ACADEMIC PRESS** *New York and London*

**111 Fifth Avenue, New York 3, New York. 17 Old Queen Street, London, S.W.1.**

# *Journal of Molecular Biology*

Published at 128s. 6d. ($18.00) per annual volume of 6 issues

Volume III Number 2, April 1961

CONTENTS



ACADEMIC PRESS INC.
LONDON NEW YORK

# *Contents*



*Printed in Great Britain by The Whitefriars Press Ltd., London and Tonbridge*